

UNIVERSITY OF ILLINOIS BULLETIN

ISSUED WEEKLY

Vol. XVIII

JANUARY 24, 1921

No. 21

(Entered as second-class matter December 11, 1912, at the post office at Urbana, Illinois, under the Act of August 24, 1912. Acceptance for mailing at the special rate of postage provided for in section 1103, Act of October 3, 1917, authorized July 31, 1918.)

BUREAU OF EDUCATIONAL RESEARCH—BULLETIN No. 5

Report of Division of Educational Tests for '19-20

BY

WALTER S. MONROE

Assistant Director

Bureau of Educational Research, University of Illinois



PRICE, 25 CENTS

PUBLISHED BY THE UNIVERSITY OF ILLINOIS
URBANA, ILLINOIS

Report of Division of Educational Tests for '19-20

BY
WALTER S. MONROE
Assistant Director
Bureau of Educational Research
University of Illinois



PRICE, 25 CENTS

PUBLISHED BY THE UNIVERSITY OF ILLINOIS
URBANA, ILLINOIS

Bulletins of the Bureau of Educational Research

B. R. BUCKINGHAM, *Editor*



Digitized by the Internet Archive
in 2023

EDITORIAL INTRODUCTION

We should like to have the reader consider this monograph as, in a certain sense, "chips from the work-shop." We hold that no organization, such as that from which this bulletin emanates, should collect from users of test materials the results which they have attained in their localities and hoard them in miserly fashion for its own purposes. Moreover, it is not so nominated in the bond. It is understood that when copies of score sheets are mailed to us we are to combine them into master score sheets and to issue tabulations which will indicate over a wider field than any school system affords the conditions disclosed by the tests in question.

Although the Bureau of Educational Research of the University of Illinois went out of the business of distributing tests last November, there had been collected up to that time a valuable body of data which was augmented during the succeeding months until it now appears to justify publication.

Aside, therefore, from chapters three and four, which deal with Monroe's Standardized Reasoning Test in Arithmetic and his Timed Sentence Spelling Test, the bulletin is devoted to presenting material which will make it possible to use a number of tests more intelligently. We do not, as Dr. Monroe says, attempt in this bulletin to give directions for administering tests or interpreting results. We are mainly concerned with what the results are. We are continually receiving questions from practical workers in the field. These questions have led us to believe that they are much interested in and perplexed by the question of standards.

Realizing this fact we have tried to make our presentation of results as complete and helpful as possible. They are presented substantially in three ways. First, we show for each test the median scores by grade. In tables devoted to this sort of data we also include the 25- and 75- percentiles. To those who understand the meaning of these latter figures the nature of the distribution out of which the medians arise will be made evident. If the 75- and 25- percentiles are far apart it means that the data are scattering. In other words that the distribution of scores spreads over a wide range.

In order that there might be no doubt about the nature of the distribution, we have in the second place presented for each test, the number of pupils attaining the indicated scores in each grade to which

the test was applied. The value of this sort of a showing is greater than the practical teacher is likely to realize upon the first inspection. Such a table may be converted into a table indicating a distribution in terms of percents by dividing each of the entires by the column totals. When the table is thus converted it becomes directly comparable with a similar table which may be computed for a school or school system. Moreover, since it is customary to give the grade medians in connection with this type of table—and the custom is followed in this bulletin—the teacher may learn from these figures the number and percent of pupils in each grade exceeding or falling short of the median of other grades. A teacher may likewise discover from such tables a number of subordinate facts concerning the test and its applicability to the grades in question—such facts as the number of zero scores or scores in the neighborhood of zero, the number of perfect or nearly perfect scores, and the nature of the distribution of the frequencies throughout.

But it is probable that the greatest usefulness of these distribution tables is of another sort. They are indispensable to those who wish to contribute toward the better standardization of tests. For example the 3000 pupils, more or less, in each grade, whose scores are shown in Table II for Monroe's Reasoning Test may reasonably be thought to be insufficient for a final standardization. This tabulation provides a form and makes a beginning for a more reliable treatment of the test in question. Any superintendent can place the pupils whom he has tested—be they few or many—in this scheme. Any bureau of research may gather scores from schools and school systems in this manner. After a little it may (and indeed it should) publish its findings in this manner to the end that more reliable standards may be secured. It is because tables of this sort are costly to print and of little direct school use that they are so seldom seen. They are frequently found after they have been converted into percentage distributions, because the latter are useful in making comparisons. But they are seldom found in mere frequency form. Yet the presentation of such tables is fundamental to cooperative effort. In our judgment every research organization ought to publish material in this form. Its high value for research purposes should be appreciated in contrast with its low value for immediate practical purposes.

On the other hand, the third form of tabulation is of most value for school uses. We are referring to the percentile tables presented in the appendix. We are convinced that when this type of material is better understood it will be much more widely used. By means of it a teacher may "place" a pupil's score among one hundred scores, arranged from highest to lowest, these one hundred scores being regarded as typical. Thus the percentile table will enable a fifth-grade teacher to state that a

pupil is (say) twentieth among one hundred typical children of his grade in speed of reading, that he is thirty-seventh in the operations of arithmetic, fiftieth in spelling, etc. If he is fiftieth in spelling we have the special case of the median, which we ordinarily arrive at from another point of view.

In using percentile tables such as those given in the appendix of this bulletin, regard must be had for the source of the tables. In its ideal form a percentile table is supposed to have been derived from a sufficient random sampling of a total "population"—e. g., from the entire fifth-grade in American schools, or from the entire number of ten-year-olds in rural schools, or from the entire number of graduates of the Chicago high schools. In ranking a child's performance one must be sure either that he belongs to the population to which the table refers or that the population of both the child and the table are indicated. Thus if a fifth-grade pupil obtains a score in composition equal to the 80-percentile for his grade, we thereby define his rank as twentieth from the top (or eightieth from the bottom) among one hundred typical fifth grade children. A pupil thus ranked has evidently done rather well compared with pupils of his own grade. Very appropriately therefore, we may wish to rank him with reference to the sixth grade. His score may perhaps equal the 60-percentile of the sixth grade. Accordingly, he would be ranked, on his performance, as fortieth from the top among one hundred typical sixth-grade children. Similarly he may rank as fiftieth (median) for the seventh grade.

We submit these percentile tables for their practical utility. They are, however, based upon a limited number of cases; and they will be somewhat modified when more scores have been made available.

From the above statements it will be clear that the chief purpose of this bulletin is to furnish an accounting of the test results which we have received. Nevertheless, we have included two chapters (III and IV) on the derivation of Monroe's reasoning tests and timed sentence tests. These accounts have been held up a long time. When, therefore, they were released, we took account of the demand that has been made for them—especially the one relative to the reasoning test—and incorporated them into this report of the Division of Educational Tests.

TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION.....	9
II. TENTATIVE GRADE NORMS	12
A. Monroe's Standardized Reasoning Tests in Arithmetic....	12
B. Buckingham's Scale for Problems in Arithmetic.....	14
C. Monroe's Diagnostic Tests in Arithmetic.....	16
D. Monroe's Standardized Silent Reading Tests.....	19
E. Charters' Diagnostic Language, and Language and Grammar Tests	25
F. Willing's Scale for Measuring Written Composition.....	26
G. Harlan's Test for Information in American History.....	28
H. Sackett's Scale in United States History	29
I. Hotz's First Year Algebra Scale	30
J. Minnick's Geometry Tests	31
K. Holley's Sentence Vocabulary Scale.....	32
L. Holley's Picture Completion Test for Primary Grades.....	34
III. THE DERIVATION OF MONROE'S STANDARDIZED REASONING TESTS IN ARITHMETIC	36
IV. MONROE'S TIMED SENTENCE SPELLING TESTS AND PUPILS' ERRORS.	48
APPENDIX: PERCENTILE SCORES.....	57
1. Monroe's Standardized Reasoning Tests in Arithmetic.	
2. Monroe's Standardized Silent Reading Tests.	
3. Charters' Diagnostic Language Tests for Grades III to VIII.	
A. Pronouns	
B. Verbs	
C. Miscellaneous A	
D. Miscellaneous B	
4. Charters' Diagnostic Language and Grammar Tests for Grades VII to VIII	
A. Pronouns	
B. Verbs	
C. Miscellaneous	
5. Willing's Scale for Measuring Written Composition	
6. Harlan's Test for Information in American History.	
7. Hotz's First Year Algebra Scale	
8. Holley's Sentence Vocabulary Scale.	

CHAPTER I

INTRODUCTION

Source of data. In distributing educational tests the Bureau of Educational Research has always supplied a duplicate class record sheet on which was printed a request that the duplicate be returned after the scores had been entered upon it. There has been no effort to follow up the purchasers of the tests in order to secure complete returns of the scores. Consequently, this report is based upon the scores voluntarily contributed. The bulk of the scores are from medium sized cities. Reports have been received from a few large cities (population of 100,000 or more) for Monroe's Standardized Silent Reading Tests but in no other case has more than one such city reported. Practically no scores were reported from rural schools except for Monroe's Standardized Silent Reading Tests. Several tests distributed by the Bureau of Educational Research are not included in this report for the reason that the number of scores reported seemed to be too small to justify any announcement of median scores which would be useful as tentative standards.

Form of the report. The distributions of scores entered upon the class record sheets were combined to form a total distribution of scores for each yearly grade. No attempt was made to keep separate the scores of the A and B sections of the yearly grades. In addition to the median scores for each grade the 25- and 75- percentile scores are also given for several of the tests. In a few cases the total distributions are given because they have a special significance. In the Appendix of this bulletin the 5-, 10-, 20-, 30-, 40-, 50-, 60-, 70-, 80-, 90-, and 95- percentile scores are given for a number of the tests. The publication of the percentile scores is prompted by the desire to make possible a more accurate interpretation of a pupil's score than merely that it is above or below the median score. Pupils belonging to any grade exhibit large individual differences. For this reason it is frequently desirable to know where the score of a pupil places him in the total distribution from which the median for his grade was calculated. The percentile scores make it possible to ascertain for any pupil his approximate position in this total distribution. For example, if a pupil's score is equal or superior to the 80- percentile score for his grade, he ranks in the upper 20 percent of all pupils to whom the test was given in that grade.

Time of testing. No attempt was made to organize the giving of the tests. Consequently, the scores on which the median and percentile scores are based represent testings all the way from September to June. This condition makes the derived scores somewhat less useful as tentative grade standards than they would be if they were based upon measures obtained at some one fixed

time during the school year. The situation is further complicated by reason of the fact that some of the schools reporting have semi-annual promotion while others have annual promotion. Some of those which have semi-annual promotion combined the A and B divisions of each grade in making their reports. In order that the median and percentile scores shall have as definite a meaning as possible we have estimated for each test the month of the school year which the median scores appear to represent. This estimate, however, must be considered only approximate.

An organized, cooperative plan which would have resulted in the tests being given on one or more fixed dates during the school year, was not attempted for two reasons. In the first place, the complete realization of such a plan was impossible because in a large number of school systems the tests were given as a part of a local plan. The Bureau of Educational Research believes it is wise to encourage this use of educational tests. When the tests are given merely as a part of a project originated by a central bureau of research, little use is likely to be made of the results by the schools giving them. Their motive is to cooperate with the central bureau, and when this has been completed there is a tendency for them to feel that all has been accomplished which may be accomplished. This results in a great waste. Information which might be of much value to the local school systems is not used. Furthermore, this practise tends to engender an attitude toward educational tests that they are merely tools of research to be used by central research bureaus and not tools which may be used by a school system, or even a teacher, in improving instruction.

A second reason for not attempting to organize the giving of the tests was that tests were distributed in all sections of the United States and in a few foreign countries. It would have been impossible to solicit the cooperation of all who gave the tests in a plan of organized testing.

The interpretation of scores by comparison with grade norms. It is not the purpose of this bulletin to give detailed suggestions concerning the use of the grade medians and percentile scores in the interpretation of the scores obtained in any school system by giving the tests. In another place¹ the writer has indicated in some detail the general procedure to be followed in interpreting scores for the purpose of improving instruction. Grade norms are also useful in interpreting the scores of pupils for the purpose of classification.²

In any interpretation of scores, either individual or group, it is necessary to bear in mind certain limitations. In the first place none of our educational tests yield scores which are absolutely accurate. The errors of measurement are large in comparison with the errors made in the measurement of physical objects. Errors larger than the difference between the median scores for suc-

¹Monroe, Walter S. "Improvement of Instruction Through the use of Educational Tests," *Journal of Educational Research*, I (February, 1920), 96-102.

²Buckingham, B.R. "Suggestions for procedure following a testing program—I, Reclassification," *Journal of Educational Research*, II (December, 1920), 787-801.

cessive school grades frequently occur although in the case of our better tests the majority of the errors are less. These errors of measurement are chance errors and for that reason tend to neutralize each other in the median and average scores of groups. Therefore, the group scores are more accurate than individual scores. However, in interpreting either type of scores one should bear in mind the possible errors of measurement which they may include.

In the second place the score which a pupil makes on any subject-matter test, such as reading, arithmetic, history, or language, depends in part upon his general intelligence. Pupils belonging to any school grade differ widely with respect to their general intelligence and consequently may be expected to differ in their achievements. For this reason some pupils belonging to a given grade should have scores above the median, while others may be expected to have scores below the median because of their differences in capacity to learn. There are also differences in the average general intelligence of pupils belonging to the same school grade. For example, the average general intelligence of the fifth-grade pupils in one school may be a year or more above that of the fifth-grade pupils in another school. It is unfair to both pupils and teachers to interpret achievement scores without recognizing the differences which may exist in the general intelligence of the pupils. To do so will frequently result in arriving at erroneous conclusions. Hence, grade standards such as are given in this report must be used with due caution.

Chapters III and IV contain reports of studies which were made by the writer during the time he was Director of the Bureau of Educational Measurements and Standards of the Kansas State Normal School, Emporia, Kansas. These reports were originally prepared for publication by that institution. Permission has been obtained to incorporate them in this bulletin. In doing this the manuscript has been only slightly revised. Chapter III gives an account of the derivation of the Monroe Standardized Reasoning Tests in Arithmetic. Chapter IV contains a description of the derivation of Monroe's Timed Sentence Spelling Tests and a report of a study of pupils' errors in spelling based upon them.

CHAPTER II

TENTATIVE GRADE NORMS

The percentile scores as well as the median scores which are given in this chapter should be used only as tentative grade standards. For several of the tests the number of scores on which these are based is so small that the standards can not be thought of as final. When other scores are added to the distributions, it is likely that different medians will be obtained. Furthermore, such standards should always be thought of as representing the average of present conditions and not as being ideal standards or what ought to be.

A. MONROE'S STANDARDIZED REASONING TESTS IN ARITHMETIC

The derivation of these tests, which consist of a series of one- and two-step problems, is described in Chapter III. For each problem two values were calculated, "correct principle value," or P, and "correct answer value," or C. These values represent the credit which is to be given for solving the problem correctly in principle and for obtaining the correct answer. Each problem is marked for correct principle. If a problem is solved correctly in principle it is further marked with reference to correct answer. A pupil does not receive credit for a correct answer if the problem was solved by the wrong principle. The directions for administering the tests provide for having the pupils mark the problem on which they are working at the end of ten minutes. In this way

TABLE I. MONROE'S STANDARDIZED REASONING TESTS IN ARITHMETIC.
FORM I. GRADE NORMS FOR APRIL TESTING

	GRADE				
	IV	V	VI	VII	VIII
CORRECT PRINCIPLE					
Number of pupils.	2932	3027	3498	2796	2472
25-percentile	6.2	12.1	10.0	13.8	11.5
Median	11.3	19.2	14.2	19.7	17.2
75-percentile	16.8	25.9	19.4	24.7	22.8
*RATE					
Number of pupils	1412	1705	1699	1717	1642
25-percentile	5.2	8.0	6.4	8.0	5.3
Median	7.8	11.2	8.7	11.2	7.5
75-percentile	8.1	15.1	12.1	14.5	10.9
CORRECT ANSWERS					
Number of pupils	2968	2996	3518	2803	2515
25-percentile	4.1	7.1	6.9	9.4	5.1
Median	7.0	11.3	10.4	13.4	9.0
75-percentile	10.7	15.5	14.0	17.4	13.0

*Sum of correct principle values of problems done correctly within ten minutes.

a rate score may be obtained. It is the sum of the "principle values" of the problems which are solved correctly in principle within ten minutes. However, the obtaining of the rate score is optional, and it was reported in only about half of the cases.

There are two forms of these tests. These forms were constructed so that they were expected to be equivalent. Experience in using them suggests that they are not equivalent, although data are lacking at this time on which a statement concerning their comparability may be based. No scores are reported for Form 2 because the returns received for this form included an insufficient number of cases.

Test I is given in Grades IV and V, Test II in Grades VI and VII, and Test III in Grade VIII. The tests were not constructed so that the scores yielded by the different tests are comparable. Therefore, direct comparisons can not be made between the fourth and fifth grade scores and between the seventh and eighth grade scores.

TABLE II. MONROE'S STANDARDIZED
REASONING TESTS IN ARITHMETIC
FORM I, CORRECT PRINCIPLE

SCORE*	GRADE				
	IV	V	VI	VII	VIII
43	3	11			
41		5			
39	7	20			
37	1	21			
35	12	89			
33	11	56			
31	26	137			56
29	25	120	47	127	94
27	40	191	93	262	63
25	80	191	131	202	171
23	89	223	135	242	214
21	124	269	248	304	233
19	130	207	280	322	214
17	161	211	306	259	217
15	248	231	328	225	237
13	260	219	470	225	203
11	298	167	425	193	201
9	267	166	374	134	133
7	294	148	276	103	154
5	304	140	191	52	121
3	230	94	123	30	87
1	185	54	49	18	51
0	137	57	22	8	23
Total	2932	3027	3498	2706	2472
Median	11.3	19.2	14.2	19.7	17.2

* In this bulletin all intervals unless otherwise noted are expressed in terms of their lower limits.

Table I gives the grade medians, 25-percentile, and 75-percentile scores for correct principle, correct answer, and rate. The distributions of scores for correct principle are given in Table II. These indicate that Test I is too difficult for a number of pupils in the fourth and fifth grades. In the construction of the tests no effort was made to include very easy problems. In fact, as is shown in Chapter III, the difficulty of a problem was not considered as a basis for selection. In none of the other grades do the zero scores amount to as much as one per cent of the total. In the seventh grade nearly five percent of the pupils made perfect scores.

REFERENCES

Willing, M. H. "The Encouragement of Individual Instruction by Means of Standardized Tests," *Journal of Educational Research*, I (March, 1920), 193-198.

Results from the Monroe Scandardized Reasoning Tests are used to illustrate how such work as the title mentions may be carried on. Suggestions for diagnosis of faults, remedial measures, etc. are given.

B. BUCKINGHAM'S SCALE FOR PROBLEMS IN ARITHMETIC.

The problems for Buckingham's scale were selected largely on the basis of difficulty. Division One is for Grades III and IV, Division Two for Grades V and VI, and Division Three for Grades VII and VIII. The problems of Division One increase by steps of approximately 0.3 P. E. from 2.7 to 5.3. The problems of Division Two increase by similar steps of difficulty from 5.5 to 7.3, and the problems of Division Three increase from 7.5 to 9.4. In scoring the test papers attention is given only to the numerical accuracy of the answers. A pupil's score is the difficulty value of the hardest problem which he answers correctly, unless he has failed on one or more previous problems. In that case, a correction is made by subtracting from the value of the hardest correctly solved problem 0.3 for each failure in Division One, or 0.2 for each failure in Division Two or Three. Thus, if a pupil solved the first six problems in Division One, his score is 4.2; but if he fails on the 4th and 5th (otherwise succeeding through the 6th), his score is 3.6—i.e., $4.2 - 2 \times 0.3$.

TABLE III. BUCKINGHAM'S SCALE FOR PROBLEMS IN ARITHMETIC. FORM I. GRADE NORMS FOR JUNE TESTING.

	GRADE					
	III	IV	V	VI	VII	VIII
No. of pupils	4181	4589	7142	5927	6632	5269
25-Percentile	3.4	4.2	5.7	5.9	7.6	7.7
Median	3.8	4.6	5.9	6.4	7.8	8.2
75-Percentile	4.3	5.2	6.3	6.8	8.3	8.7

Although the three divisions of the scale were constructed so that it was expected that the scores obtained from the different divisions would be com-

parable, the grade medians given in Table III clearly indicate that the scores are not comparable. The increase in the median scores from the third grade to the fourth grade is 0.8. The increase from the fourth grade to the fifth grade is 1.3. A similar variation is found in the differences between the subsequent grades. Therefore, the scores obtained by the different divisions of the scale are not comparable. The reason for this is that the pupils taking Division Two or Division Three do not have an opportunity to do the problems of the lower divisions. If they did, a number of them would fail to do all of them correctly. Thus, they would receive a score lower than that which they receive when taking only the higher divisions.

TABLE IV. BUCKINGHAM'S SCALE FOR
PROBLEMS IN ARITHMETIC. FORM I
GRADE DISTRIBUTIONS FOR
JUNE TESTING

SCORE	GRADE					
	III	IV	V	VI	VII	VIII
9.0					328	699
8.5				6	782	1084
8.0		1	4	11	1349	1290
7.5			14	13	2931	1740
7.0			240	775	58	44
6.5		2	1012	1886		
6.0		6	1540	1504		
5.5	2	21	3663	1569		
5.0	131	1069	106	57		
4.5	490	1474	14	12		
4.0	815	863				
3.5	1305	798				
3.0	967	255				
2.5	298	75				
0	173	25	549	94	1184	412
Total	4181	4589	7142	5927	6632	5269
Median	3.8	4.6	5.9	6.4	7.8	8.2

In Table IV the total distributions are given. Evidently a division of the scale higher or lower than that designed for the grade has been used in a few cases. The distributions are significant in that they show that the divisions of the scale are too difficult for the respective grades. The percent of pupils making zero scores in the third, fifth, seventh, and eighth grades is so large that the scale as now published must be considered unsatisfactory for these grades. This condition could be remedied in the case of Division Two and Division Three by giving the next lower division to the pupils who make zero scores. In the case of Division One, the scale will have to be extended downward by adding less difficult problems.

REFERENCES

First Annual Report, Bureau of Educational Research, University of Illinois, pp. 21-22. These pages contain a very brief suggestion of what was done along the line of this scale that seemed to justify its construction, also a short description of the scale.

Buckingham, B.R. "Notes on the Derivation of Scales in School Subjects, with Special Application to Arithmetic," *Fifteenth Yearbook of National Society for the Study of Education*, Part I, pp. 23-40.

This presents a report of a series of problems which was given to a number of school children in New York and other cities. The results are given and discussed, especially with reference to locating the problems on a scale. Although this scale is not the one now in use, it is similar to it.

C. MONROE'S DIAGNOSTIC TESTS IN ARITHMETIC.

Monroe's Diagnostic Tests in Arithmetic consist of four parts. Part I includes Tests 1 to 6. Part II includes Tests 7 to 11. These tests involve only integers. Part III includes Tests 12 to 16, which consist of examples involving common fractions. Part IV includes Tests 17 to 21, which consist of examples involving multiplication and division of decimal fractions. Tables V and VI

TABLE V. MONROE'S DIAGNOSTIC TESTS IN ARITHMETIC. GRADE
MEDIAN FOR APRIL TESTING. RATE (NUMBER
OF EXAMPLES ATTEMPTED)

	GRADE				
	IV	V	VI	VII	VIII
PART I					
(Approximate number of pupils)	900	480	590	600	600
Test 1	7.2	11.6	13.3	12.6	14.0
Test 2	4.1	7.2	9.3	8.6	9.2
Test 3	3.3	5.0	5.8	5.7	7.2
Test 4	2.0	3.2	4.0	4.7	5.7
Test 5	3.9	4.8	5.6	5.7	6.2
Test 6	1.7	2.7	3.1	3.0	4.0
PART II					
(Approximate number of pupils)	380	760	610	520	460
Test 7	3.8	4.2	5.5	5.3	6.3
Test 8	3.0	4.1	5.5	6.1	6.6
Test 9	4.8	5.9	8.2	8.6	9.8
Test 10	2.8	3.2	5.3	5.3	6.7
Test 11	1.6	2.2	2.2	2.9	3.7
PART III					
(Approximate number of pupils)		370	1000	580	560
Test 12		5.8	7.6	8.6	9.4
Test 13		4.5	5.4	6.0	6.0
Test 14		5.2	7.1	8.1	8.7
Test 15		6.2	7.1	7.7	8.1
Test 16		5.7	7.4	8.0	9.1
PART IV					
(Approximate number of pupils)			440	900	660
Test 17			3.6	3.5	4.5
Test 18			11.9	11.5	12.9
Test 19			5.8	4.5	5.3
Test 20			12.5	11.1	13.5
Test 21			5.1	4.3	4.8

give the median scores for these tests in terms of rate (number of examples attempted) and accuracy (percent of examples done correctly). In order to simplify the administration of these tests the plan of scoring has been changed so that the pupil is now given only one score, the number of examples right. In Table VII tentative grade norms are given in terms of this score.

In the interest of economy, both of cost of the tests and time required for their administration, most of the tests of this series were made so short that there is a lack of discrimination between pupils. For example, the increase in the number of examples attempted from grade to grade is frequently less than one example. The shortness of the tests also makes the errors of measurement relatively large.

This group of tests was designed for diagnostic purposes, i. e., it was intended to measure separately the abilities of pupils to do the important types of

TABLE VI. MONROE'S DIAGNOSTIC TESTS IN ARITHMETIC. GRADE
MEDIAN FOR APRIL TESTING. ACCURACY (PERCENT
OF EXAMPLES CORRECT)

	GRADE				
	IV	V	VI	VII	VIII
PART I					
Approximate number of pupils	900	480	590	600	600
Test 1	100	100	100	100	100
Test 2	66.6	86.8	100	100	100
Test 3	56.5	72.3	80.8	82.0	87.6
Test 4	28.0	55.1	71.9	79.8	85.4
Test 5	52.5	61.7	66.9	67.9	75.1
Test 6	22.4	49.5	64.0	77.5	100
PART II					
Approximate number of pupils	380	760	610	520	460
Test 7	63.2	65.1	75.9	76.3	81.6
Test 8	30.4	52.9	66.9	79.8	78.5
Test 9	75.0	86.3	91.2	93.1	100
Test 10	35.2	58.4	72.1	72.3	81.9
Test 11	22.4	35.5	53.4	65.0	68.2
PART III					
Approximate number of pupils		370	1000	580	560
Test 12		35.5	32.0	33.6	49.0
Test 13		38.0	29.2	36.0	53.9
Test 14		57.5	70.3	79.6	86.0
Test 15		37.5	30.0	33.6	45.5
Test 16		38.5	36.8	59.1	70.2
PART IV					
Approximate number of pupils			440	900	660
Test 17			37.6	36.4	53.2
Test 18			100	100	100
Test 19			39.6	47.0	61.7
Test 20			100	100	100
Test 21			35.6	44.0	51.3

examples in the field of the operations of arithmetic. A weighted sum of a pupil's scores on such a group of tests would yield a general measure of his ability in this field.³

TABLE VII. MONROE'S DIAGNOSTIC TESTS IN ARITHMETIC. GRADE MEDIAN FOR APRIL TESTING. NUMBER OF EXAMPLES CORRECT

	GRADE				
	IV	V	VI	VII	VIII
PART I					
Approximate number of pupils	900	480	590	600	600
Test 1	7.2	11.6	13.3	12.6	14.0
Test 2	2.8	6.2	9.3	8.6	9.2
Test 3	1.9	3.6	4.7	4.7	6.4
Test 4	.6	1.7	2.9	3.6	4.8
Test 5	2.0	3.0	3.7	3.8	4.4
Test 6	.4	1.3	2.0	2.3	4.0
PART II					
Approximate number of pupils	380	760	610	520	460
Test 7	2.4	2.7	4.2	4.0	5.1
Test 8	.9	2.2	3.7	4.8	5.2
Test 9	3.6	5.1	7.5	8.0	9.8
Test 10	1.0	1.8	3.8	3.8	5.5
Test 11	.4	.8	1.2	1.8	2.4
PART III					
Approximate number of pupils		370	1000	580	560
Test 12		2.0	2.4	2.9	4.6
Test 13		1.7	1.6	2.2	3.2
Test 14		3.0	5.0	6.5	7.4
Test 15		2.3	2.1	2.6	3.7
Test 16		2.2	2.7	4.7	6.4
PART IV					
Approximate number of pupils			440	900	660
Test 17			1.4	1.3	2.4
Test 18			11.9	11.5	12.9
Test 19			2.4	2.1	3.3
Test 20			12.5	11.1	13.5
Test 21			1.8	1.9	2.5

REFERENCES

Finley, G. W., *A Comparative Study of Three Diagnostic Arithmetic Tests*. Colorado State Teachers College Bulletin, Series XX, No. 4.

This reports a study made of the Cleveland Survey Tests, The Woody Arithmetic Scales and Monroe's Diagnostic Tests in Arithmetic. The tests were given on six successive days to some 60 eighth grade pupils. The scores made are given in detail, compared with each other and with scores obtained elsewhere.

Monroe, W. S., "A Series of Diagnostic Tests in Arithmetic," *Elementary School Journal*, XIX, (April, 1919), 585-607.

This article discusses the types of examples in the four fundamental operations, the question of "one dimensional" vs. "two dimensional" tests, and thus establishes the theoretical bases of the tests presented. The series is described, a distribution of scores made thereon is analyzed, and the value of using such tests pointed out.

³See the group of tests on the operations of arithmetic included in the Illinois Examination.

Uhl, W. L., "The Use of Standardized Materials in Arithmetic for Diagnosing Pupils' Methods of Work," *Elementary School Journal*, XVIII, (November, 1917), 215-218. This article contains no reference to the Monroe Tests, but describes an experiment in diagnosis similar to that made possible by their use. Both finding specific faults and remedying them is considered briefly.

D. MONROE'S STANDARDIZED SILENT READING TESTS.

Monroe's Standardized Silent Reading Tests have been used so widely that a detailed description here is unnecessary. Each test consists of several exercises, each of which has been assigned a rate value and a comprehension value. The rate value is based upon the number of words in the exercise and the comprehension value is based upon the rate and accuracy with which pupils were found to be able to do the exercise. Test I is for Grades III, IV, and V, Test II is for Grades VI, VII, and VIII, and Test III is for the high school. There are three forms of Tests I and II. There are only two forms of Test III.

The different forms of these tests were constructed so that they were expected to be equivalent. The use of the forms, however, indicates that they are not equivalent. In order to study the degree of equivalence of the three forms, copies of the different forms were arranged in alternate order before the test papers were distributed to the pupils. This plan results in the first, fourth, seventh, tenth, etc. pupil having a copy of Form 1. The second, fifth, eighth, eleventh, etc. pupil would have a copy of Form 2. The third, sixth, ninth, twelfth, etc. pupil would have a copy of Form 3. By this plan each form of the test is given to similar samples of the school population.

Test I was given to approximately 775 pupils and Test II was given to approximately 645. The numbers of pupils taking the different forms in each grade differed slightly. This is an accidental result of the way in which the test papers were arranged. The average and the standard deviation have been calculated for each distribution of scores. In general the pupils made higher scores on Forms 2 and 3 than they did on Form 1. The standard deviations are also unequal. This suggests that the exercises of the different forms of the tests make somewhat irregular scales.

The formula for reducing the scores obtained from one scale to equivalent scores on another scale is as follows:

$$S_1 = \frac{\sigma_1}{\sigma_2} S_2 + \left(Av_1 - \frac{\sigma_1}{\sigma_2} Av_2 \right)$$

In this formula S_1 is the equivalent score in Form 1 and S_2 the obtained score in Form 2. Av_1 refers to the average of the scores obtained from Form 1, Av_2 refers to the average of the scores obtained from Form 2. σ_1 is the standard deviation of the distribution of the Form 1 scores and σ_2 is the standard deviation of the distribution of the Form 2 scores. This formula is based upon the usual assumption that the deviations from the average are equal when expressed in terms of the standard deviation of the distribution; in other words that

$$\frac{S_1 - Av_1}{\sigma_1} = \frac{S_2 - Av_2}{\sigma_2}$$

When this equation is solved for S_1 we obtain the formula as given above. Since the scores on Form 1 are in general smaller than the scores on the other two forms it was decided to reduce both the Form 2 and the Form 3 scores to the equivalence of Form 1 scores. The application of the above formula involves the determination of the numerical value of the ratio of $\frac{\sigma_1}{\sigma_2}$ by which the Form 2 score is to be multiplied and the determination of the numerical equivalent of the constant terms of the formula, (i.e., of the expression in parentheses). This latter numerical equivalent may be plus or minus. When it is positive it is to be added and when negative it is to be subtracted.

TABLE VIII. MONROE'S STANDARDIZED SILENT READING TESTS
AVERAGES, STANDARD DEVIATIONS AND CORRECTION NUMBERS TO REDUCE
FORM 2 AND FORM 3 TO THE EQUIVALENCE OF FORM 1.

COMPREHENSION	FORM	No. PUPILS	Av.	STANDARD DEVIATION	CORRECTION NUMBERS	
					MULTIPLIER	CONSTANT TERM
TEST I	1	775	11.4	6.5	-	-
	2	777	12.0	5.9	1.100	-1.8
	3	769	12.0	6.2	1.040	-1.1
TEST II	1	649	22.2	9.5	-	-
	2	643	23.8	9.0	.944	-2.6
	3	639	25.2	9.6	.885	-1.1
RATE						
TEST I	1	775	63.8	24.0	-	-
	2	777	72.7	27.5	.870	+6
	3	769	80.4	29.1	.820	-2.1
TEST II	1	649	97.2	29.5	-	-
	2	643	117.7	32.9	.896	-8.3
	3	639	104.3	22.7	1.300	-38.4

$$\text{Formula: } S_1 = \frac{\sigma_1}{\sigma_2} S_2 + (Av_1 - \frac{\sigma_1}{\sigma_2} Av_2)$$

In the above table the first correction number or multiplier is $\frac{\sigma_1}{\sigma_2}$ and the second or constant term is

$$(\text{Av}_1 - \frac{\sigma_1}{\sigma_2} \text{Av}_2).$$

In Table VIII, the number of pupils, the average score and the standard deviation is given for each form of each test. In the last two columns the multiplier and the constant term in the above formula are given for Form 2 and Form 3. These can be used in reducing scores obtained from Form 2 or Form 3 to the basis of Form 1. In securing data for these determinations, each test was given in each of the three grades for which it is intended. Except in the eighth grade the number of pupils in the different grades was approximately the same. The correction numbers were calculated for each grade separately. Since they were found to be approximately the same it was decided to combine the scores from the different grades and compute a single set of correction numbers for each test.

The grade medians calculated from the distributions of the scores yielded by the different forms furnished additional information concerning the degree of their equivalence. This information is not in complete agreement with that obtained by the study described. Although it is less accurate it deserves some consideration in the formulation of a set of rules for translating the scores obtained from one form to the basis of another form. In Table IX-A grade medians for all forms are given, and in Table IX-B correction numbers which may be used in reducing scores from Form 2 or Form 3 to Form 1. The correction numbers are based primarily on the results of the study just described but some weight was given to the information furnished by the tabulations of

TABLE IX-A. MONROE'S STANDARDIZED SILENT READING TESTS. GRADE MEDIAN FOR JANUARY AND JUNE TESTING, BASED UPON 130,000 SCORES

	GRADE									
	III	IV	V	VI	VII	VIII	IX	X	XI	XII
FORM 1										
Rate										
January	52	70	87	90	100	106	83	85	90	96
June	60	79	94	96	104	108	86	87	94	100
Comprehension										
January	6.8	12.7	17.8	18.5	22.8	26.0	23.0	25.4	27.2	30.0
June	9.3	15.3	20.8	21.0	24.5	27.3	24.0	26.0	28.6	32.0
FORM 2										
Rate										
January	63	77	98	116	130	133	84	90	98	104
June	70	88	106	124	132	136	86	92	101	109
Comprehension										
January	8.3	13.3	17.2	18.1	26.0	28.2	25.4	28.0	31.0	33.1
June	10.6	15.6	20.5	20.8	27.3	29.4	26.6	29.4	32.2	34.5
FORM 3										
Rate										
January	78	92	97	101	109	111				
June	85	95	104	106	111	114				
Comprehension										
January	9.3	14.8	18.4	22.2	26.5	29.8				
June	11.9	16.8	21.5	24.4	28.2	30.5				

TABLE IX-B. APPROXIMATE CORRECTIONS BY WHICH TO MULTIPLY FORM 2 AND FORM 3 SCORES TO REDUCE TO THE BASIS OF FORM 1 SCORES

RATE			COMPREHENSION		
Test I	Test II	Test III	Test I	Test II	Test III
Form 2 .88	.80	.94	.95	.93	.90
Form 3 .78	.93		.94	.86	

the scores obtained from the different forms. This is the explanation of some apparent inconsistencies in the reductions to the basis of Form 1.

It should be noted that the scores of the different tests in this series are not comparable. This is to be expected in the case of the rate scores but in the case of the comprehension scores an effort was made to have the different tests yield comparable scores. This attempt was not successful.

The grade distributions which are not published here, show that the tests are too short for the time allowed. In order to secure accurate measures of the abilities of the best readers it will be necessary either to lengthen the test or to shorten the time allowed. The wide spread use of these tests has revealed other defects. Instead of attempting to remedy these defects in the present series it was decided to derive an entirely new series. These have been issued under the title of "Monroe's Standardized Silent Reading Tests, Revised." Three forms of Tests I and II are now available. They were originally published as a part of the Illinois Examination but are now printed separately.

In Tables X-A and X-B we have assembled a miscellaneous collection of grade medians. These are published because a number of requests have been received for just this type of information.

REFERENCES

Monroe, W. S., "Monroe's Standardized Silent Reading Tests," *Journal of Educational Psychology*, IX, (June, 1918), 303-312.

The derivation of the tests more or less based upon the Kansas Silent Reading Test plan, is briefly sketched. The investigation of weighting and timing is outlined, a sample of the tests is given and some few data concerning results from pupils.

Barnes, Harold, "Reorganization of Classes Based on the Monroe Silent Reading Tests," *University of Pennsylvania Bulletin*, vol. XX, No. 1, 119-123.

This article recounts the use made of these tests in the elementary grades of Girard College. Not only are the scores presented, but also the resulting organization upon the basis of ability as shown on the tests is described.

Kelly, F. J., "Kansas Silent Reading Tests," *Journal of Educational Psychology*, VII, (February, 1916), 63-80.

The author of these tests, which were the forerunners of Monroe's Standardized Silent Reading Tests, gives a brief statement of the construction, administration, and use of the tests, following it with a more detailed statement of results secured in nineteen Kansas cities.

Lloyd, S. M. and Gray, C. T., "Reading in a Texas City, Diagnosis and Remedy," *University of Texas Bulletin*, No. 1853.

This bulletin gives an account of a study of the reading situation in Austin. The Monroe tests were given in grades 3-7. Results obtained are analyzed at considerable length, measures to improve the situation are discussed, and improvement after a period of special emphasis on reading is shown.

Pressey, S. L. and L. W., "The Relative Value of Rate and Comprehension Scores in Monroe's Silent Reading Test, as Measures of Reading Ability," *School and Society*, (June 19, 1920), 747-49.

In a brief discussion of the above subject, the writers present results of correlating teachers' estimates of reading ability with rate and comprehension scores, also the latter with each other. They conclude that comprehension scores may tell us all the tests can about children's ability in reading.

TABLE X-A. MONROE'S STANDARDIZED SILENT READING TESTS
GRADE MEDIAN, RATE

GRADE											
	III	IV	V	VI	VII	VIII	IX	X	XI	XII	
Form 1 Large Cities (May) Medium Sized Cities (May)	57.0 43.6	70.5 65.0	88.1 78.5	91.6 88.7	110.0 94.0	106.2 106.4	82.5	85.6	86.6	102.6	
Form 2 Large Cities (May) Medium Sized Cities (May) Wisconsin Cities (February) Denver (October) Denver (May) Pittsburgh (May)	64.1 60.1 69.0 31.0 53.7 65.0	77.0 74.5 76.0 59.0 76.0 78.0	97.1 97.9 103.0 76.0 95.0 93.0	114.9 113.2 115.0 81.0 115.0 105.0	133.7 130.1 133.0 98.0 127.5 117.0	138.3 135.9 133.0 98.0 133.2 133.0	87.4 89.0	88.0 89.0	89.2 106.0	101.6 106.0	
Form 3 Large Cities (May) Medium Sized Cities (May)	81.4 80.1	101.6 93.3	111.0 105.4	102.7 100.9	113.5 108.8	119.0 112.8					
Form NOT SPECIFIED Pennsylvania Cities and Boroughs (April)*	63.0	78.0	89.0	105.0	115.0	136.0					

*See *Seventh Annual Schoolmen's Week Proceedings*, University of Pennsylvania Bulletin, Vol. XXI, No. 1, p. 122.

TABLE X-B. MONROE'S STANDARDIZED SILENT READING TESTS
GRADE MEDIANS, COMPREHENSION

	GRADE											
	III	IV	V	VI	VII	VIII	IX	X	XI	XII		
FORM 1												
Large Cities (May)	7.9	12.1	18.9	18.6	24.6	27.3						
Medium Sized Cities (May)	4.9	12.0	14.6	18.8	20.6	24.1	24.0	25.8	28.5	33.3		
FORM 2												
Large Cities (May)	9.2	13.7	17.8	23.5	27.7	30.3						
Medium Sized Cities (May)	7.7	13.1	17.9	21.7	25.3	29.3						
Wisconsin Cities (February)	8.9	13.4	18.6	21.4	25.0	27.9	23.8	27.2	29.6	31.6		
Denver (October)	2.0	7.5	11.0	16.0	19.0	21.0	26.5	29.8	32.9	37.5		
Denver (May)	8.0	12.5	16.0	22.0	24.5	30.0						
Pittsburgh (May)	8.5	13.5	16.0	18.5	23.5	28.5						
FORM 3												
Large Cities (May)	10.9	14.5	19.5	23.2	26.6	32.5						
Medium Sized Cities (May)	9.6	15.2	19.0	22.4	26.5	31.0						
FORM NOT SPECIFIED												
Pennsylvania Cities and Boroughs (April)*	8.3	12.1	15.7	18.0	22.9	27.6						

*See *Seventh Annual Schoolmen's Week Proceedings*, University of Pennsylvania Bulletin Vol. XXI, No. 1, p. 122.

E. CHARTERS DIAGNOSTIC LANGUAGE TESTS, AND DIAGNOSTIC LANGUAGE AND GRAMMAR TESTS.

There are two groups of these tests: (1) The Diagnostic Language Tests, designed for Grades III to VIII inclusive, which include, Pronouns, Verbs, (formerly Verbs A), Miscellaneous A (formerly Miscellaneous), Miscellaneous B (formerly Verbs B); (2) The Language and Grammar Tests, designed for Grades VII and VIII, which include Pronouns, Verbs (formerly Verbs A) and Miscellaneous A (formerly Miscellaneous). The Language Tests consist of a number of sentences most of which are grammatically incorrect. If a sentence is correct the pupil makes a cross on the dotted line below the sentence. If the sentence is not right the pupil is required to put the correct words on the dotted line below it. In the Language and Grammar Tests the pupil is required in addition to write the rule on which the correction is based. The pupil's score is the number of exercises which he does correctly. Since the sentences which make up the tests were selected as representative of the errors which pupils make, a pupil's performance on the tests gives a diagnosis of his abilities in the field of these tests.

There are two forms of these tests. The second form, however, was not published until September, 1920. Consequently, the scores reported in this bulletin are based on Form 1. Although the two forms were constructed so that Form 2 might be expected to be equivalent to Form 1, there is available at this time no information concerning the degree of their equivalence.

TABLE XI A. GRADE NORMS FOR CHARTERS' DIAGNOSTIC LANGUAGE TESTS. MARCH TESTING

	GRADES					
	III	IV	V	VI	VII	VIII
*MISCELLANEOUS A						
Number of Pupils	386	669	668	845	758	494
25-percentile	4.0	5.8	8.1	11.8	14.0	16.6
Median	6.7	9.3	11.6	16.5	18.9	22.3
75-percentile	13.3	13.6	16.0	21.7	24.4	27.1
†MISCELLANEOUS B						
Number of Pupils	230	430	307	475	412	294
25-percentile	3.0	10.6	15.7	19.8	23.5	28.7
Median	7.9	17.8	22.0	27.3	29.4	32.0
75-percentile	14.8	24.5	27.6	32.4	33.7	36.8
**VERBS						
Number of pupils	365	403	373	478	539	638
25-percentile	7.3	12.9	17.2	19.0	22.7	28.6
Median	12.6	17.7	22.6	24.3	27.7	32.8
75-percentile	18.8	22.7	28.4	29.3	31.9	36.1
PRONOUNS						
Number of pupils	787	864	895	1344	1566	1253
25-percentile	8.9	11.1	14.2	17.0	19.6	23.1
Median	13.6	15.1	18.5	21.4	24.5	29.0
75-percentile	19.8	20.3	22.6	25.7	29.5	34.0

* Formerly Miscellaneous

† Formerly Verbs B

** Formerly Verbs A

TABLE XI-B. GRADE NORMS FOR
CHARTERS' DIAGNOSTIC LANG-
UAGE AND GRAMMAR TESTS
MARCH TESTING

	GRADES	
	VII	VIII
MISCELLANEOUS		
Number of pupils	332	362
25-percentile	2.9	6.1
Median	6.3	11.9
75-percentile	11.7	18.7
VERBS		
Number of pupils	434	497
25-percentile	2.8	6.9
Median	7.8	14.0
75-percentile	22.9	24.1
PRONOUNS		
Number of pupils	332	362
25-percentile	4.4	8.5
Median	8.0	17.1
75-percentile	16.7	26.0

REFERENCES

Charters, W. W., "Minimum Essentials in Elementary Language and Grammar," *Sixteenth Yearbook of the National Society for the Study of Education*. Part I, 85-110.

This article gives a brief account of a number of studies of language and grammar errors made by school children, with tables of results. These studies were the basis of the content of Charters' tests.

Sixth Conference on Educational Measurements. Bulletin of the Extension Division, Indiana University, Vol. V, No. 1, pp. 6-12 and 13-24.

These two discussions by Charters give a rather general discussion leading up to a brief account of the development and form of the tests, followed by some suggestions as to their use.

Charters, W. W., "Constructing a Language and Grammar Scale," *Journal of Educational Research*, I (April, 1920), 249-257.

The tests herein considered are a revision of those referred to above. The writer gives a short description of their derivation, use, scoring, etc. The question of weighting is discussed and the reason for its elimination given.

F. WILLING'S SCALE FOR MEASURING WRITTEN COMPOSITION

The Willing Scale for Measuring Written Composition differs from other composition scales in that an attempt is made to secure separate measures of "form value" and "story value." The "form value" of a pupil's composition is based upon his errors in grammar, punctuation, capitalization, and spelling. In order to make the scores in form value comparable, the number of errors which the pupil makes is multiplied by 100 and divided by the number of words in his composition. The quotient is the number of errors per hundred words. The "story value" of a pupil's composition is its value when errors of grammar punctuation, capitalization, and spelling are neglected. This value is measured by means of the scale.

TABLE XII. GRADE NORMS FOR WILLING'S SCALE FOR MEASURING WRITTEN COMPOSITION, MARCH TESTING

	GRADE					
	III	IV	V	VI	VII	VIII
APPROX. NO. OF PUPILS	325	580	705	695	570	130
STORY VALUE						
25-percentile	30.5	43.4	55.7	60.9	65.9	65.3
Median	41.5	58.7	68.1	74.0	76.6	79.0
75-percentile	54.8	74.5	78.8	85.2	86.6	86.2
FORM VALUE (Errors per 100 Words)						
25-percentile	11.7	6.2	3.6	3.2	2.6	2.3
Median	18.5	10.7	6.8	5.8	4.4	4.4
75-percentile	26.0	17.3	10.9	9.8	6.4	7.0

The grade distributions given in Tables XIII and XIV indicate that the scale needs to be extended at both ends. It does not contain steps low enough in story value to provide adequate measures for many compositions contributed by pupils in the third and fourth grades. Neither does it provide adequate measures for the best compositions in grades beyond the fourth. For practical purposes these limitations are not serious because when a pupil's composition is as poor as 20 on this scale the pupil needs special attention. When a pupil writes a composition as good as 90 on this scale it is likely that special instruction is superfluous. In addition it may be pointed out that the median score of a class is probably not affected by these limitations of the scale.

TABLE XIII. WILLING'S SCALE FOR MEASURING WRITTEN COMPOSITION. GRADE DISTRIBUTIONS FOR MARCH TESTING

Story Value Score*	GRADE					
	IV	V	VI	VII	VIII	IX
90	7	28	52	81	95	14
80	6	67	105	177	141	48
70	8	91	168	147	150	25
60	28	91	146	123	107	21
50	60	98	110	75	41	12
40	60	90	77	49	31	4
30	75	66	39	32	5	4
20	76	48	8	8	2	1
Total	320	579	705	692	572	129
Median	41.5	58.7	68.1	74.0	76.6	79.0

*These intervals are expressed in terms of their mid-points.

REFERENCES

Willing, M. H., "The Measurement of Written Composition in Grades IV to VIII," *English Journal*, VII (March, 1915), 193-202.

The writer explains and outlines the measurement of written composition especially in connection with the Denver and Grand Rapids surveys. The method of constructing the scale, its use, scoring, results obtained, etc. are discussed, and the scale reproduced.

The *Denver Survey*, 1916, Part II, pp. 59-63, and the *Grand Rapids Survey*, 1916, pp. 85-105, give accounts of the use of this scale. The latter contains rather complete tables and graphs of pupil achievement, and comparisons of results with those obtained in Denver

TABLE XIV. WILLING'S SCALE FOR MEASURING
WRITTEN COMPOSITION. GRADE DISTRI-
BUTIONS FOR MARCH TESTING

Form Value*	GRADE					
	IV	V	VI	VII	VIII	IX
30	50	14	3	5	2	
27	24	17	1	4	3	
24	22	27	8	4	2	
21	32	29	9	4	1	
18	43	38	19	18	4	
15	29	51	45	17	15	2
12	42	56	52	48	28	4
9	34	120	105	96	49	12
6	32	79	149	140	131	21
3	14	98	170	199	200	49
0	5	41	141	158	138	41
Total	327	570	702	693	573	129
Median	18.5	10.7	6.8	5.8	4.5	4.4

* Errors per 100 words.

G. HARLAN'S TEST FOR INFORMATION IN AMERICAN HISTORY

This test consists of ten exercises in the field of American History and is designed for use in the seventh and eighth grades. Each exercise consists of two or more parts. The maximum score which a pupil may receive is 100.

TABLE XV. GRADE NORMS FOR
HARLAN'S TEST OF INFORMATION
IN AMERICAN HISTORY. MAY
TESTING

	GRADE	
	VII	VIII
Number of pupils	1109	1691
25-percentile	30.1	45.7
Median	43.9	68.2
75-percentile	57.3	83.3

In Table XVI the distributions of scores for the seventh and eighth grades are given. These distributions are of interest because of the very great

individual differences which they suggest. It is possible that the apparent differences are due in a considerable measure to the errors of measurement. Since there is only one form of the test no measure of reliability is available.

TABLE XVI. HARLAN'S TEST OF
INFORMATION IN AMERICAN
HISTORY. GARDE DISTRI-
BUTIONS FOR MAY
TESTING

SCORE	GRADE	
	VII	VIII
96	6	66
91	3	118
86	13	140
81	18	187
76	35	136
71	35	136
66	54	112
61	65	92
56	65	100
51	99	95
46	113	79
41	115	98
36	100	90
31	93	79
26	94	61
21	79	42
16	63	35
11	37	13
6	19	10
0	3	2
TOTAL	1109	1691
MEDIAN	43.9	68.2

REFERENCES

Harlan, Chas. L., "Educational Measurement in The Field of History," *Journal of Educational Research*, II (December, 1920), 849-853.

The writer follows a short discussion of tests in the "content" subjects with a brief description of his test and its use in nine cities. The requirements he deems essential to a good test are listed as a basis of his test.

Griffith, G. L., "Harlan's American History Tests in the New Trier Township Schools," *School Review* (November, 1920), 697-708.

The first half of this article is devoted to a general discussion of history and history testing. This is followed by a description of the test and the results of its use in the eighth grade of this township. Data are given for each of the single exercises of the test.

H. SACKETT'S SCALE IN UNITED STATES HISTORY

This scale, arranged by L. W. Sackett, was originally devised by Bell and McCollum. It consists of seven tests which appear to have been intended

for use in secondary schools and colleges. The medians given in Table XVII are for the eighth grade. The number of scores is such that it is doubtful if the median scores have much value for use as tentative standards.

TABLE XVII. GRADE NORMS FOR SACKETT'S SCALE IN UNITED STATES HISTORY. MAY TESTING. EIGHTH GRADE

	TESTS						
	I	II	III	IV	V	VI	VII
Number of pupils	111	101	107	92	93	78	85
25-percentile	62.6	50.8	69.2	37.0	44.7	7.5	9.6
Median	118.7	146.2	115.0	125.0	86.5	46.6	96.5
75-percentile	192.5	273.9	183.1	287.5	193.7	138.1	195.5

REFERENCES

Bell, J. C. and McCollum, D. F. "A Study of the Attainments of Pupils in United States History," *Journal of Educational Psychology*, VIII (May, 1917), 257-74.

The writers follow a discussion of historical ability with an account of the use of test material in various schools from Grade V through the senior year of the University of Texas. The results secured are analyzed. The test questions used were in general similar in kind to those of Sackett's Scale in Ancient History, although based upon United States History.

Sackett, L. W. "A Scale in Ancient History." *Journal of Educational Psychology*, VIII (May, 1917), 284-93.

The test questions are given with a brief statement of their source, use, and scoring. Results are given from almost 1000 papers, and the relative difficulty of the questions computed.

Sackett, L. W. "A Scale in United States History," *Journal of Educational Psychology*, X (September, 1919), 345-348.

The writer tells of the development of this scale out of the data furnished by Bell and McCollum's work referred to above. The determination of the relative difficulty of the parts is given considerable space.

I. HOTZ'S FIRST YEAR ALGEBRA SCALE

This scale consists of five separate scales: (1) Addition and subtraction; (2) Multiplication and division; (3) Equation and formulae; (4) Problems; (5) Graphs. Each sub-scale consists of exercises arranged in order of increasing difficulty.

REFERENCES

Holz, H. G. *First Year Algebra Scales*, Teachers College, Columbia University, Contributions to Education No. 90.

The writer gives a history of the derivation of these scales, a complete reproduction of them, and a discussion of their administration and use. The statistical working out of the scales is treated fully for one of them, the procedure for all being the same.

Cawl, F. R. "Practical Uses of an Algebra Standard Scale," *School and Society* (July, 1919), 89-91.

The results of testing a class in a large private school are here presented. The matter of correlation with English, French, and Latin is considered. A short interpretation of results is given, with suggestions as to the value of using such a scale.

TABLE XVIII. GRADE NORMS FOR
HOTZ'S FIRST YEAR ALGEBRA
SCALES. MAY TESTING

	GRADE	
	IX	X
ADDITION AND SUBTRACTION		
Number of pupils	561	390
25-percentile	5.2	5.8
Median	6.9	7.3
75-percentile	9.1	8.7
MULTIPLICATION AND DIVISION		
Number of Pupils	570	388
25-percentile	5.7	5.9
Median	7.2	7.4
75-percentile	8.4	8.7
EQUATIONS AND FORMULAS		
Number of Pupils	478	385
25-percentile	6.2	6.7
Median	7.7	7.9
75-percentile	9.7	9.1
PROBLEMS		
Number of Pupils	566	394
25-percentile	4.5	3.9
Median	6.4	5.0
75-percentile	8.6	6.3
GRAPHS		
Number of Pupils	121	413
25-percentile	5.2	4.1
Median	6.2	5.0
75-percentile	7.0	6.0

J. MINNICK'S GEOMETRY TESTS

This series of tests is based on the assumption that the demonstration of a geometrical theorem involves the following abilities: Test A, the ability to draw the figure. Test B, the ability to state the hypothesis and conclusion. Test C, the ability to recall the facts concerning the figure. Test D, the ability to select and organize facts so as to produce the proof. Test E, the ability to draw auxiliary lines. The series includes one test for each of these abilities. No report is made for Test E. These tests are unique in that they provide for both positive scores and negative scores. The positive score is the percent of the necessary elements of the proof given correctly by the pupil. The negative score is the number of incorrect and unnecessary elements.

REFERENCES

Minnick, J. H. *An Investigation of Certain Abilities Fundamental to the Study of Geometry.* University of Pennsylvania.

This monograph gives a synopsis of methods and results used in deriving the tests, followed by a more detailed statement. The latter includes a reproduction of the tests, tables giving data secured from testing, statistical methods of weighting exercises, suggestions as to use, etc.

TABLE XIX. GRADE NORMS FOR MINNICK'S GEOMETRY TESTS

	Positive Scores Grade		Negative Scores Grade	
	X	XI	X	XI
TEST A (ability to draw accurate figures for theorems)				
Number of pupils	126	66	126	60
25-percentile	53.3	43.8	2.4	1.1
Median	63.0	58.0	4.1	2.6
75-percentile	67.2	69.2	6.6	5.4
TEST B (Ability to state hypothesis and conclusion in terms of given figure.)				
Number of Pupils	167	66	167	66
25-percentile	55.2	55.5	1.1	1.0
Median	69.6	67.1	2.3	2.0
75-percentile	81.3	83.6	3.9	3.9
TEST C (Ability to recall known facts about figures when one or more are given).				
Number of Pupils	154	65	154	63
25-percentile	52.2	55.6	1.9	1.4
Median	64.1	64.7	3.8	3.9
75-percentile	77.2	77.9	7.1	5.7
TEST D (Ability to organize and select facts to produce a proof).				
Number of Pupils	155	68	155	54
25-percentile	68.0	75.0	.8	.8
Median	85.5	89.2	1.6	1.6
75-percentile	92.9	98.3	2.3	3.2

Minnick, J. H. "A Scale for Measuring Pupil's Ability to Demonstrate Geometrical Theorems," *School Review*, (Feb., 1919), 101-109.

A brief account of the construction of a scale to measure one definite geometric ability is given. The scores made upon the first selection of exercises, the resultant weighting and then the selection of those best suited to make up a scale are briefly treated. The exercises chosen are reproduced.

Minnick, J. H. "Certain Abilities Fundamental to the Study of Geometry," *Journal of Educational Psychology*, (Feb., 1918), 83-90.

Four abilities requisite to formal geometrical demonstration are listed. Their relation to teaching, development by teaching, and diagnosis by tests are discussed. The tests used were those of the author. Correlations with teachers marks are given.

K. HOLLEY'S SENTENCE VOCABULARY SCALE

This scale consists of a number of exercises of the following type:

1. *Impolite* people are.....kindly.....brave.....young.....ill-bred.
2. A man is *afloat* in a.....mine.....tower.....boat.....hospital.

The pupil is asked to underline the word which makes the truest sentence. These exercises are arranged in order of increasing difficulty, and a pupil's score is found by subtracting one-third of the number of errors from the number correct. An abbreviated form of this scale has been incorporated in the Illinois General Intelligence Scale. The scale was constructed to provide a suitable means of ascertaining the general intelligence of groups of children. The measure which it yields is not sufficiently accurate to be used as an index of the general intelligence of individual pupils. The scale is also recommended as

an instrument for measuring the vocabulary of pupils. The total distributions given in Table XXI indicate that this scale is too difficult for pupils in the third and fourth grades.

TABLE XX. GRADE NORMS FOR HOLLEY'S SENTENCE VOCABULARY SCALE
APRIL TESTING

	GRADE									
	III	IV	V	VI	VII	VIII	IX	X	XI	XII
Number of pupils	406	520	465	450	1188	1047	253	223	155	108
25-percentile	8.4	16.7	25.3	33.4	32.3	40.1	40.9	50.1	52.4	54.5
Median	16.6	25.1	33.0	42.8	41.9	47.7	49.0	56.0	59.9	62.7
75-percentile	28.5	33.6	39.9	51.5	49.7	55.8	57.1	63.5	67.9	70.1

REFERENCES

Terman, L. M. and Childs, H. G. "A Tentative Revision and Extension of the Binet-Simon Measuring Scale of Intelligence," *Journal of Educational Psychology*, (April, 1912), 205-208.

The basis of Holley's Sentence Vocabulary Scales is the Stanford Revision, 100 word Vocabulary Test, the construction of which is here described. Tentative standards of achievement are also given.

Holley, C. E. *Mental Tests for School Use*. Bureau of Educational Research, University of Illinois, Bulletin No. 4 pp. 86-91.

This bulletin gives an account of a comparative study of six group intelligence scales, of which the above was one, based on data from the school system of Champaign, Illinois. A brief account of the origin of the Sentence Vocabulary Scales is included (p. 30).

Branson, E. P. "An Experiment in Arranging High-School Sections on the Basis of General Ability," *Journal of Educational Research*, (Jan., 1921), 53-56.

At Long Beach, California, this scale was given to two groups of high-school entrants who had recently taken, and been grouped by the Otis Group Intelligence Scale. At the end of the term the test was repeated. A comparison by groups of the scores at the two periods, and correlations with the Otis Scale, are given.

TABLE XXI. HOLLEY'S SENTENCE VOCABULARY SCALE. GRADE DISTRIBUTIONS FOR APRIL TESTING

SCORE	GRADE									
	III	IV	V	VI	VII	VIII	IX	X	XI	XII
90					6					
80					5			1	2	7
70					3			21	27	21
60		1	5	50	75	126	32	56	48	36
50	31	4	20	74	199	321	79	85	52	31
40	18	44	89	141	379	341	78	45	23	13
30	42	128	170	110	290	153	50	15	3	
20	71	170	137	59	123	73	5			
10	124	130	30	13	74	29	1			
0	120	43	14	3	34	3				
TOTAL	406	520	465	450	1188	1047	253	223	155	108
MEDIAN	16.7	25.1	33.0	42.8	41.9	47.8	49.0	56.0	59.9	62.7

L. HOLLEY'S PICTURE COMPLETION TEST FOR PRIMARY GRADES

This test, as its name suggests, consists of a number of pictures which are incomplete. The pupil is expected to add the part which is missing. It was designed as an instrument for measuring the general intelligence of young children. The total distributions as given in Table XXIII indicate that it is not a good instrument for this purpose. The distributions exhibit unusually

TABLE XXII. GRADE NORMS FOR HOLLEY'S PICTURE COMPLETION TEST FOR PRIMARY GRADES. JANUARY TESTING

	GRADE				
	Kinder- garten	I	II	III	IV
Number of pupils	75	1438	1233	327	167
25-percentile	1.8	4.4	7.9	9.9	8.8
Median	5.3	7.8	11.5	13.5	12.3
75-percentile	7.8	12.2	15.1	16.5	15.0

TABLE XXIII. HOLLEY'S PICTURE COMPLETION TEST FOR PRIMARY GRADES. GRADE DISTRIBUTIONS FOR JANUARY TESTING

SCORE	GRADE			
	I	II	III	IV
20	34	31	9	4
19	23	43	13	4
18	33	46	25	9
17	36	59	20	12
16	42	64	27	3
15	52	72	34	10
14	47	76	22	15
13	54	88	26	14
12	48	98	24	17
11	82	77	19	14
10	72	92	24	9
9	81	84	24	12
8	91	89	26	9
7	103	75	11	11
6	122	59	9	13
5	94	50	8	8
4	112	50	4	2
3	108	29	1	1
2	86	21	1	
1	77	28		
0	41	2		
TOTAL	1438	1233	327	167
MEDIAN	7.8	11.5	13.5	12.3

high variability. This is much greater than is exhibited by other tests when applied to children in these grades. The median scores given in Table XXII give further indications of the inadequacy of this test, particularly in the grades above the first.

REFERENCES

Holley, C. E. *Mental Tests for School Use*. Bureau of Educational Research, University of Illinois, Bulletin No. 4, pp. 86-91.

A general discussion of tests of this type is followed by an account of the testing from which this test came. This was done in Champaign, Illinois. Results are merely outlined.

CHAPTER III

THE DERIVATION OF MONROE'S STANDARDIZED REASONING TESTS IN ARITHMETIC⁴

The process of problem solving. "Reasoning" as it occurs in the solving of an arithmetical problem involves these steps: (1) A careful reading of the problem including the association of correct arithmetical meanings with the "technical" terms used in stating the problem. (2) Recall of facts and principles suggested by the problem and required for its solution. (3) Formulation of a hypothesis or plan of solution using as data the results of the first two steps. (4) Verification of this plan of solution. This process of reasoning is usually followed by the calculations outlined in the plan of solution. This additional step, however, is not a part of the reasoning process.

Two kinds of words are used in stating arithmetical problems: (1) The descriptive words give the setting of the problem. Only in an indirect way do these affect the solution. (2) The "technical terms" of an arithmetical problem consist of those words and phrases which define quantities and quantitative relationships. Every problem involves at least three quantities, two given and the third to be found. These quantities are related in a definite way. For example, the sum of the two quantities given equals the third, or the third is the quotient of one divided by the other. In problems involving two or more steps there are more than three quantities and the relationships are more complex. However, in every case there are words or phrases which either directly or indirectly tell what these relationships are, and, consequently, what operations must be performed to obtain the desired answer.

This principle may be illustrated by the following problems: "What are the average daily earnings of a boy who receives \$0.88, \$0.25, \$1.15, \$0.75, \$0.50, and \$0.60 in one week?"

The phrase "average daily earnings" names the quantity to be found and also specifies its relationship with the given quantities. The "average" is the quotient of the sum of the several amounts divided by the number of items. A knowledge of this definite meaning of "average" is necessary if one is formulating a rational plan of solving the problem. If the phrase "average daily" was omitted we would have an entirely different problem.

"How many square yards of linoleum will be required to cover a floor 16 feet by 12 feet?"

"How many square yards" names the third quantity in this problem and in connection with "15 feet by 12 feet" specifies the relations which exist

⁴A number of considerations on which the derivation of these tests is based are contained in an article by the writer in *School and Society*, Volume VIII, pages 295 and 424. Sample copies of these tests may be obtained from the Public School Publishing Company, Bloomington, Illinois.

between the quantities. This third quantity is the product of the dimensions divided by nine.⁵ In this case the number of square feet in a square yard must be remembered and also the principle that the area of a rectangle (i.e., a figure whose dimensions are given as in the problem) is the product of the length by the width.

In many cases when the first two steps of the reasoning process have been completed satisfactorily, the formulation of the plan of solution (the next step in the reasoning process) involves little uncertainty. In fact it is essentially mechanical. This is the case in these illustrations. In the case of very simple problems, or very familiar problems, the reasoning process is usually short-circuited so that there is no explicit association of meaning with the technical terms nor recall of principles. The problem as a whole or some feature of it serves as a cue for the direct association of the plan of solution. In such cases there is strictly speaking, no reflective thinking or reasoning, and the mental process involved is much the same as that which occurs in the operations of arithmetic. The solution of the problem has become automatic.

The nature of a reasoning test in arithmetic. A reasoning test in arithmetic is essentially a test of careful reading in a limited field to answer specific questions. In this reading, technical vocabulary is fundamental. The pupil gives evidence of his degree of comprehension by his plan of solution. The correctness of the numerical answer to the problem depends upon the accuracy of the pupil's calculations and the recall of denominate number facts as well as upon the plan of solution. The *plan*, or principle, of the solution and not the *accuracy* of the numerical answer is, therefore, the measure of the pupil's ability to reason in arithmetic. Thus in describing a pupil's performance on a reasoning test, errors in the recall of facts and in calculation should be disregarded. For the problems which are solved correctly in principle a score based on correct answers may be used as a crude measure of the pupil's ability to perform the operations of arithmetic.

In order that a pupil's score on a reasoning test may be indicative of his ability to solve arithmetical problems in general, the problems must be carefully selected with reference to content (vocabulary). The ideal reasoning test would be one that included all of the technical terms but this is not possible because the vocabulary of arithmetical problems is extremely varied and voluminous. In another⁶ place the writer has reproduced 28 different forms of statement which were found in the examination of eight text-books for the problem, "Given, \$7.50, paid for silk, and price per yard \$1.50, to find the number of yards purchased." This condition makes it necessary to select a few problems which will be representative in respect to content, in order to have a test of usable length.

⁵An alternative solution is to reduce each dimension to yards before finding the area.

⁶Monroe, Walter S. *Measuring the Results of Teaching*, Houghton Mifflin Company (1918), 163.

Method of selecting problems on basis of content. In the case of the series of tests described in this report the representative problems were selected by the following method. The one- and two-step problems appearing in eight widely used texts were classified according to the operation or operations they called for. This gave in one group all the problems requiring only addition, in another those requiring only subtraction and so on. The problems in each of these groups were further classified by the writer on the basis of the technical terms used. This was found to be difficult because of the great variety of these terms. Since the classification represents the judgment of only one person, it cannot be considered final in any sense.

The general plan of classification may be illustrated by the types of division problems. In the list on the following pages only those types are given which included problems found in five or more of the eight texts examined. A large number of additional types included problems from less than a majority of the texts. A descriptive statement of the type is followed by a limited number of illustrative problems. It will be noted that for a single type these problems are not identical in vocabulary but it was the judgment of the writer that they were sufficiently similar to justify grouping them together. It has been assumed that the terms used are essentially synonymous. This hypothesis is, of course, subject to experimental verification. Unfortunately, this is lacking at this time. However, the resulting list of type problems is more representative of the vocabulary of arithmetical problems as they occur in our texts than any other available list.

DESCRIPTION OF TYPES AND ILLUSTRATIVE PROBLEMS:

1. Given a whole to find number of parts of a given size, including to find the number of acres to produce a given yield.

A baker used three-fifths lbs. of flour to a loaf of bread. How many loaves could he make from a barrel?

When the average yield per acre is 25 bushels how many acres will yield 925 bushels.

How many lengths three-fourths yds. long can be cut from 15 yds. of goods?

How many hens can be properly accommodated in a pen containing 51 square feet, if each hen requires 6 square feet?

964 marbles are distributed equally among a certain number of boys. Each boy has 82 marbles. How many boys are there?

At $7\frac{1}{2}$ gallons to the cubic foot, how many cubic feet will 3000 gallons of oil occupy?

Oats weigh 32 lbs. to a bushel. How many bushels are there in a load weighing 1344 lbs?

2. Given cost and price to find the number of articles purchased. This includes wages when question is how many days, weeks, etc. to earn a given amount.

At 16 cents per pound, how many pounds of steak does a woman get if the amount of the purchase is 80 cents.

3. The reverse of No. 1. Given whole and number of parts to find size of each part.

Three boys buy a rowboat for twelve dollars and seventy-five cents, sharing the expense equally. Find how much each boy has to pay.

If 54 marbles are divided equally among 6 boys, how many marbles will each receive?

In 28 days a hotel used 361 lbs. of butter. How many pounds did it use a day?

4. The reverse of No. 2.

A farmer paid thirty-three dollars and a half for 4 bushels of seed wheat. How much did he pay for a bushel?

The bill for 58 tons of copper amounted to 612 dollars. What was the price per ton?

A fowl weighing 4 and one half lbs. sells for \$1.00. What is the price per pound?

A man's wages amounted to 46 dollars for 9 and one-fifth day's work. How much did he receive per day?

A man works 8 hours a day for 4 dollars and 80 cents. How much does he receive for each hour's work?

5. Given the price for a given denomination to find the price at a lower denomination.

A boy bought a dozen oranges at the rate of 15 cents a dozen. What did they cost him apiece?

When milk is 10 cents a quart, how much is a pint worth?

6. Given the whole and the number of parts to find the average (rate, price, yield, etc.)

A farmer raised 500 bushels of wheat on a field of 40 acres. What was the average yield per acre?

A fast train runs from Chicago to a station 356.4 mi. distant in exactly 9 hours. What is the average rate of the train?

A drover paid \$1125 for cows, what was the average price if he bought 25?

A mill employs 600 hands and has a weekly pay roll of \$2,000. What is the average weekly wage for each employee?

7. The whole and the rate are given. The question is asked by, "How long?"

If a horse eats three-eighths bu. of oats a day, how long will 6 bus. last?

How long will it take to earn 28 dollars at \$1.75 a day?

8. Given distance and rate to find how long.

At 25 miles an hour, how long will it take an automobile to go 160 miles?

9. Given distance and number of units of time to find rate.

In 3.2 hours a man walks 12.32 mi. How far does he walk in one hour?

Find the rate of speed per hour made by an airship traveling 218.05 miles in 3.5 hour.

10. A fractional part of a whole is given to find the whole.

If, when 18 and three eighths mi. of track are laid, one third of the road is completed, how long is the road?

I sold a bicycle for 18 dollars. This was three sevenths of what I paid for it. How much did I pay for it?

11. A percent of the whole is given to find the whole.

If 33 and one third percent of a man's loss is 300 dollars, how much does he lose?

A girl spent 25 cents which was $12\frac{1}{2}$ percent of her monthly allowance, how much was her allowance?

A clerk had his weekly wages increased 3 dollars, or 16 and two thirds percent. What were his wages before this increase?

12. Given the amount of gain or profit and percent of gain or profit to find the cost or selling price.

A hardware merchant makes a profit of 25 percent or 32 cents on saws. Find the cost.

A farmer sold his horse at a gain of 30 dollars, or 25 percent. Find the cost.

13. Given the commission and percent of commission to find amount sold.

Five percent commission on a certain amount of money was 684.20 dollars. What was the amount?

14. Given two numbers to find what percent one is of the other.

If 1000 lbs. of potatoes contain 180 lbs. of starch, what percent of potatoes is starch?

If a man saves 187.50 dollars out of his salary of 1250 dollars, what percent does he save?

The boys in the Marshall school won 5 of the 8 games of hockey. What percent?

In his examination in arithmetic a boy had 10 problems out of twelve right. His grade was what percent?

15. Given two numbers to find what part one is to the other.

The Jackson basket-ball team won 35 out of 56 games. What part did it win?

A man spends for rent 360 dollars out of an income of 1500 dollars. What part of his income is spent this way?

16. Given the amount of investment, or principle, and the income or interest to find rate.

Mrs. Lynch received 24 dollars a year interest on 400 dollars loaned Mrs. Burnet. What is the rate?

17. Given an amount in one denomination to reduce to a higher.

An aviator reaches a height of 11,474 feet. Express this height in miles.

A milk dealer sells 302 qts. of cream. Express this as gallons and quarts.

In digging out a cellar 8260 cubic feet of earth were removed. At 27 cubic feet to the cubic yard, how many cubic yds. were removed?

18. Given the value or face of a policy and premium to find the rate.

Find the rate, given the face of the policy as 1500 dollars and premium 15 dollars.

A fire insurance company charged 20 dollars for insuring an automobile for 1000 dollars. What was the rate of insurance.

19. Given the premium and rate of insurance to find face of policy.

A man paid 50 dollars for insuring a house, the rate being 2 and $\frac{1}{2}$ percent. What was the face of the policy?

Table XXIV gives the frequency of occurrence of each type in each of the eight texts examined. This table is to be read as follows: 10 problems classified as belonging to Type 1 which were found in Text 1; 30 such problems were found in Text 2; 20 in Text 3; 34 in Text 4, etc. The total number of problems classified under Type 1 is 147.

The variations in the frequency of the occurrence of problems belonging to a single type are worthy of notice. Some types have a high frequency in certain texts while in other texts their frequency is low and in many cases they do not occur at all. This means that different authors have tended to use different vocabularies.

TABLE XXIV. FREQUENCY OF OCCURRENCE OF TYPES OF PROBLEMS IN DIVISION IN EIGHT TEXTS

TYPE NUMBER	TEXT								Total
	1	2	3	4	5	6	7	8	
1	10	30	20	34	8	13	13	19	147
2	28	14	123	11	-	75	49	27	327
3	-	7	6	2	4	11	3	15	48
4	11	24	24	2	19	20	14	17	131
5	2	4	2	-	3	1	-	1	13
6	5	10	4	9	3	13	4	52	100
7	-	2	2	3	2	1	1	2	13
8	1	1	-	4	-	4	1	-	11
9	2	7	3	-	1	1	2	12	28
10	-	1	13	2	2	13	3	2	36
11	6	19	13	3	-	4	11	16	72
12	1	-	2	-	1	1	3	-	8
13	-	1	1	2	4	1	-	1	10
14	9	42	12	41	3	14	39	30	190
15	-	12	5	-	1	-	18	4	40
16	1	2	1	1	1	8	1	8	23
17	6	1	1	1	-	2	2	-	13
18	8	-	4	-	1	-	3	7	23
19	9	-	3	1	-	-	1	1	15

Space does not permit the reproduction of similar tables for addition, subtraction, multiplication and the classes of two-step problems. In Table XXV a summary of the frequencies of the occurrence of the several types is given. This table is read as follows: In the case of problems requiring only addition, three types occurred in all eight texts, one occurred in seven and two in six texts. The total number of types occurring in five or more of the texts is six. The total number of problems classified in these six types is 464. The total number of problems is 622.

The reader should bear in mind that no attempt was made in this classification to determine what problems pupils should be asked to solve. The

problems have been taken as they occurred in the texts. In effecting the classification, no consideration was given to the question of whether the problem was practical. In fact the purpose was not to obtain a list of practical problems but to secure a list of the forms of statement or language which had been used in the one- and two-step problems by the authors of widely used texts. Many of the technical terms of arithmetic are used (probably must be used) whether the problems are practical or not.

Experimental selection of problems. In order to secure data for the construction of a series of reasoning tests in arithmetic, about 300 problems were selected out of the total number examined and classified. Out of this number 156 problems were chosen for an experimental series of tests. In making the selections for this purpose the writer considered, in addition to the classification described above, the social importance of the problems. Thus a few types of problems which occur in a majority of the texts and have a high total frequency, were not represented. This introduces an additional subjective factor but in view of the emphasis which is being placed upon the social importance of the subject matter, the writer believes it is better to exercise judgment in this instance rather than to follow blindly statistics based upon the content of our present texts, particularly when it is obviously impossible to include representative problems of all types within a single series of tests of suitable length for classroom use.

TABLE XXV. FREQUENCY OF TYPES OF PROBLEMS

OPERATION	NUMBER OF TYPES OCCURRING IN				TOTAL No. OF TYPES	FREQUENCY OF PROBLEMS CLASSIFIED	FREQUENCY OF ALL PROBLEMS
	8 texts	7 texts	6 texts	5 texts			
+	3	1	2	0	6	464	622
-	1	2	4	0	7	211	456
×	5	5	2	3	16	1641	1938
÷	5	6	3	5	19	1248	1610
+ -	1	2	2		5	199	346
+×	4	2	3	4	13	472	718
+÷	1	0	1	3	5	127	299
-×	0	0	1	7	8	166	559
-÷	1	0	3		4	114	413
×	3	1	2	4	10	429	581
÷	1	4	2	6	13	704	1234
++							7
÷÷							70
TOTAL	25	23	25	32	105	5775	8853

In constructing the experimental series, Test I was designed for grades four and five, Test II for grades six and seven, and Test III for grade eight. Some such division is necessary because certain social situations from which

problems are taken are not studied until the later grades although the mathematical relationships are very simple. Pupils cannot be expected to solve such problems until they are acquainted with the social situations. For this reason all problems involving percentage were placed in Test III. No consideration was given to the relative difficulty of the problems in making this division except that no problems requiring common fractions were placed in Test I and for the most part decimal fractions were confined to Test III.

Each test consisted of sixteen problems printed on a four page folder with space so that the pupil could do all of his work upon the test paper. The test papers showed that unless the pupil made errors and did his work over or used an elaborate method, ample space was provided except in a very few cases. The directions for administering the preliminary tests were essentially the same as those which now accompany the tests.

A number of cities were invited to cooperate by giving the tests between April 1 and 15, 1918. Fourteen cities responded, nine in Kansas, and one city in each of the following states: Illinois, Ohio, Michigan, New York, and Pennsylvania. Usable returns were received from 12,859 pupils.

When the record sheets and the test papers were returned to the writer it was found that the directions for marking the papers were not sufficiently complete and explicit. Consequently, there was a lack of uniformity in the marking. In order to insure uniformity the writer, assisted by two clerks rescored the papers. Whenever an unusual or questionable solution was found a record was made and all similar solutions were marked in the same way. In this way a high degree of uniformity in the marking of the papers was secured. Space does not permit a detailed statement of the plan of scoring of the solution of each problem but the general plan may be indicated.

The solution of a problem was considered correct in *principle* if the pupil's work showed that he had based his solution upon the relationships which exist between the quantities of the problem. For example, in the problem, "If a man has \$275 in the bank and draws out \$70, how much has he left in the bank?", there are *three* quantities: \$275, \$70, and the amount "left in the bank." These are related so that the difference between \$275 and \$70 must equal the amount left in the bank. A solution of the problem based upon this relationship must involve the subtraction of \$70 from \$275, or the finding of a number which added to \$70 will make \$275.

In the problem, "A house rents for \$35 a month. This is how much a year?", the three quantities are \$35, 12, or the number of months in a year, and the amount for a year. The relation is that the product of \$35 and 12 equals the amount of rent for a year. A solution based upon this relationship would usually be one in which \$35 was multiplied by 12. In a few cases the pupil had set down \$35 twelve times and added. This solution was counted as correct in principle because it was considered that the pupil had recognized

the relationship which existed between the quantities of the problem. Incidentally it should be noted that although such a solution was counted as being correct in scoring the papers of the test, a teacher should not encourage it. In fact the writer believes it should be discouraged, except possibly when the pupil is learning the idea of multiplication, because the method is not efficient. It requires more time and there are more opportunities for error in the mechanical work.

In the case of the above problem, if a multiplier other than 12 was used the solution was counted as correct in principle because it was considered that correct recall of denominate number facts was not a part of the reasoning. In a few cases 35 was multiplied by itself. This was marked incorrect in principle.

Although the pupils were directed to do all work upon the test papers a few gave only the answer in the case of certain problems. They had either solved the problem mentally or on another sheet of paper. An arbitrary rule was adopted. If the answer was correct the problem was marked correct in principle and answer. If the answer was wrong it was marked incorrect in both principle and answer.

An answer was not marked correct unless the solution of the problem was correct in principle and the answer was numerically correct and in its lowest terms if it contained a fraction. It was not required that the answer be labeled with its denomination.

Weighting the problems. For each problem three records were secured: (1) Number of pupils attempting the problem. (2) Number of solutions correct in principle. (3) Number of correct answers. From these facts the percent of solutions correct in principle and the percent of those solved according to the right principle which had also correct answers were calculated. These percents were translated into *sigma* values. The former being designated as the "P" value of the problem and the latter as the "C" value. In doing this it was assumed that the ability to solve problems was distributed normally and included between $+2.5$ *sigma* and -2.5 *sigma*. The tables given in Rugg's "Statistical Methods as Applied to Education" were used. The values were calculated to two decimal places but in order to simplify the computation of scores they were expressed in terms of the nearest integer in the tests as now published.

In the case of those problems which were solved by the pupils in two successive grades, the average inter-grade interval was found for each group of problems by taking the average of the differences of the *sigma* values of the problems of the test. This inter-grade interval was added to the values of the problems for the upper of the two grades to reduce them to the basis of the lower grade. The average of the two values was taken as the final value of the problem.

An attempt was made to reduce the *sigma* values to a common zero point, and thus secure comparable scores, by having a limited number of problems from Test I appear in Test II and also a limited number of problems from Test II appear in Test III. It happened that some of the problems chosen showed inversion and for this reason it was deemed advisable not to attempt to reduce the values to a common zero. Thus the scores obtained from the different tests of the series are not comparable.⁷

Construction of the final tests. Out of the 156 problems included in the preliminary test, 90 which belonged to types occurring in five or more of the eight texts examined, were selected for the final tests. Since in the selection on the basis of content there was no effort to include problems which exhibited wide range of difficulty, there was no attempt to construct a difficulty scale. In fact it is the judgment of the writer that the educational objectives implied by such a scale in the field of problem-solving are open to serious criticism. In our schools we should endeavor to instruct pupils to solve problems because they are socially worth while rather than because they exhibit a certain degree of difficulty. The purpose here is to construct a group of tests containing problems that are representative of the language in which problems are stated in our representative text books and which appear to be satisfactory for testing purposes.

The final tests consist of 15 problems each. Test I is for grades four and five, Test II for grades six and seven, and Test III for grade eight. There are two forms of each test. In selecting the 90 problems for these tests those were rejected which were commented on unfavorably by those who gave the preliminary tests. Also those problems were rejected which were found to be particularly confusing to pupils. The arrangement of the order of the problems in a test was made without reference to their difficulty values. An attempt was made to secure as high degree of variation in the operations required as possible. In the two forms of each test the corresponding problems are approximately equal in difficulty, and so far as possible, the two forms were made equivalent in other respects.⁸

⁷The method of weighting is open to criticism. It is used in an attempt to give more credit for doing a difficult problem than for doing an easy one. It is not at all certain that such a plan gives the most truthful indication of a pupil's ability. Some recent studies have shown that unweighted scores correlate very highly with the weighted scores obtained by this method. Therefore, it is likely that the tests would have been nearly as accurate measuring instruments without any determination of weights.

⁸No determination of the reliability or validity of these tests was made as a part of the original derivation. Neither is it possible to make a report on these questions at this time. Some work which was done on the question of reliability indicated that the tests were less reliable than tests in the operations of arithmetic and in silent reading. This appeared to be due to the fact that frequently pupils are unable to do certain problems because of a peculiar course of study.

Analysis of errors made by pupils on preliminary test. In the preliminary testing the following six problems were given to 100 fifth grade pupils in one city. The results of an analysis of the test papers are given in Table XXVI.

1. Mrs. Black received \$2 a yd. for broadcloth. She sold 78 yds. How much did she receive?

2. At the store a towel roller costs 35c. George made one for his mother. He used 12c worth of lumber, 2c of hardware, and 3c worth of shellac. Find how much George saved his mother.

3. A Kansas farmer bought 80 acres of cheap land for \$240. Oil being found on his farm he sold the land for \$60,000. What was his profit?

4. A car contains 72,060 lbs. of wheat. How much is it worth at 87c a bushel?

5. A field is 20 rds. long and 12 rds. wide. How many rods of fence are needed to enclose it?

6. What are the average daily earnings of a boy who received 88 cents, 25 cents, \$1.15, 75 cents, 50 cents, and 60 cents in one week?

TABLE XXVI. RESULTS OF ANALYZING THE ERRORS OF 100 FIFTH GRADE PUPILS

	PROBLEM						Total
	1	2	3	4	5	6	
Number of pupils attempting	100	100	92	52	94	94	532
Errors in reasoning	8	21	39	38	33	67	206
Errors in fundamentals	1	7	26	16	2	35	87
Omissions and errors in copying	2	0	2	3	0	4	11
Errors in decimals	13	0	16	25	0	14	68

Two significant facts are shown in this table. First, a majority of the errors (55 percent) are in reasoning. More than one-third of the attempts (39 percent) resulted in faulty reasoning. Second, 41 percent of the errors in calculation were in placing the decimal point. This second fact becomes more significant when we note that these errors occurred in problems involving only United States money and that the first and third problems which produced 29 of the 68 errors do not really involve decimal fractions. In these two problems the error consisted in pointing off the answer when it should not have been done.

The wide spread use of the Courtis Standard Research Tests, Series B and other tests upon fundamentals has resulted in increased attention to the fundamental operations with integers. There should be no decrease in the emphasis upon this phase of arithmetic for one out of six pupils made errors in the simple calculations required in these problems but the greatest source of error and therefore the greatest need is increased attention to the mental processes involved in reasoning as it occurs in solving arithmetical problems. The necessity for doing this becomes more obvious when we examine the nature

of the errors in reasoning. In problem 2, 13 of the 21 errors in reasoning were due to adding all of the terms; in problem 4, 33 out of 38 of the errors in reasoning were due to multiplying 72,060 by 87 without attempting to reduce the pounds to bushels; in problem 6, 58 of the 67 pupils who reasoned incorrectly simply added the terms. Each of these errors may be ascribed to inaccurate or incomplete reading of the problem, or the first step in the rational solution of a problem.

In tabulating the scores of the preliminary test the variation in achievement of different classes was particularly noticeable. It was evident that some teachers were teaching their pupils to solve problems while others, frequently within the same school system, were not doing so. It is also significant that a number of teachers consistently marked as correct, solutions which were clearly incorrect. This may have been accidental on the part of the teacher but this is very doubtful. A striking illustration was furnished by the problem: "A baker used $3/5$ lbs. of flour to a loaf of bread. How many loaves could he make from a barrel (196 lbs.) of flour?" The correct solution requires that 196 be *divided* by $3/5$ which gives an answer of $326 \frac{2}{3}$ loaves. In several instances teachers marked as correct all papers in a class in which 196 was multiplied by $3/5$. This latter solution gives an answer of $117 \frac{3}{5}$ loaves. The fact that teachers made such errors as this indicates that they are not familiar with solving reasoning problems. Perhaps this is one source of the poor records made by the pupils.

CHAPTER IV

MONROE'S TIMED SENTENCE SPELLING TESTS AND PUPIL'S ERRORS

How ability in spelling should be measured. In measuring ability in spelling by having a pupil spell words which are dictated in lists it is clear that the conditions under which the pupil spells the words are not the conditions under which he spells the words which he uses in writing themes, letters, and other school exercises. As a result we probably fail to obtain a measure of his "true spelling ability."

If the test words are embedded in sentences and the sentences written from dictation we approach more nearly normal spelling conditions because the pupil is writing connected words which have meaning. A still closer approximation appears to be secured by dictating the sentences at approximately the rate at which the pupil is accustomed to write. By thus causing the pupil to write at approximately his normal rate of writing, he does not have time to study over the spelling of words, and as a result we secure a record of spelling which is largely automatic. Under such conditions a pupil's attention is centered primarily upon writing and not upon his spelling. Of course, these conditions are not those under which the pupil normally spells words. The writing from dictation may be an unusual exercise for the pupil. Some pupils will be accustomed to write more slowly than the rate of dictation. This may tend to confuse them. To what extent these and possibly other factors prevent our obtaining a record of the "true spelling ability" of the pupil by a timed-dictation test we do not know. It appears, however, that a *Timed Sentence Spelling Test* is likely to yield a more valid measure of spelling ability than a list of words dictated separately.

The construction of Monroe's Timed Sentence Spelling Tests.

In order to make easily available a timed sentence spelling test, the writer constructed a series of such tests, using test words chosen from appropriate columns of Ayres' Scale for Measuring the Ability in Spelling and basing the rate of dictation upon the measurements of the rate of handwriting of over six thousand Kansas school children. In order that the scores might have a high degree of reliability as measures of the spelling ability of individual pupils, fifty test words were used in each test. According to one study⁹ the probable error of an individual score for a test of fifty words is less than 1.00 when the score is expressed as the percent of words spelled correctly. For a class of twenty-five or more pupils the probable error of the class score would be 0.2.

⁹Otis, A. S. "The reliability of spelling scales involving a 'deviation formula' for correlation." *School and Society*, 4, (November 11, 1916), 716-22. This study deals with the reliability of tests consisting of isolated words and it is possible that the results might not apply to a timed sentence spelling test.

For grades III and IV the test words were taken from Column M of Ayres' Scale for Measuring Ability in Spelling. For grades V and VI they were taken from Column Q and for Grades VII and VIII and the high school they were taken from Coulmans S, T, and U. The test for the fifth grade is reproduced. The sentences are to be dictated when the second hand of the watch reaches the position indicated in the left-hand margin.

In this test no test words come at the end of the sentences. Thus, the pupil who writes slowly will be much less likely to make low scores because he does not have time to complete writing the sentences. It should also be noted that all other words found in these sentences are easier to spell as shown by the Ayres Scale. It was thought advisable to allow time for dictating the sentences in addition to the actual writing. For this reason, the rate of writing for each school grade was increased by ten percent, sixty-six seconds instead of sixty being allowed for the number of letters which pupils commonly write in sixty seconds.

Tentative grade Norms. This series of timed sentence spelling tests was given in sixteen Kansas cities in April and May, 1917. During the school year of 1919-20 scores were reported from a number of cities. The grade medians for these two groups of cities and the norms given by Ayres are given in Table XXVII. In comparing the successive grades it must be remembered that the same test words were not used for all grades. One list of test words was used for grades III and IV, another for grades V and VI and still another for grades VII and VIII and for the high school.

The fact that the median scores in Table XXVII are materially below Ayres' norms indicates that a different type of spelling ability has been measured by the timed spelling sentence test than that measured by Ayres in constructing his scale. (Ayres had the words dictated in lists). This fact becomes more apparent when it is recalled that many of the cities which gave these tests had used Ayres' Scale as a minimum course of study as well as a source of test words. Thus, had the test words been dictated in lists it is likely that the median scores would have been materially above Ayres' norms.¹⁰

MONROE'S TIMED SENTENCE SPELLING TEST ARRANGED FOR THE FIFTH GRADE.

Seconds

- 60 The *president* gave *important information* to the men.
- 48 The *women* were *present* at the time.
- 19 The *entire region* was burned over.
- 49 The *gentlemen* declare the *result* was printed.
- 30 Suppose a *special attempt* is made.

¹⁰It is possible that the difference between the median scores and Ayres' norms may be due to factors other than the measurement of different types of spelling ability. Many of the pupils probably were not accustomed to writing from dictation and all were not accustomed to writing at the rate at which these tests were dictated. It is possible that these unusual conditions may have been operated to materially lower the scores of a number of pupils or even of most pupils.

- 60 The *final debates* were held.
 24 The *factory employs* forty men.
 51 *Sometimes* the *connection* is not made.
 24 I *enclose a written statement* with the book.
 3 *Prompt action* is needed.
 25 It was a *wonderful surprise* to all.
 55 The *addition* to the *property* was begun.
 31 *Remember, Saturday* is the day.
 57 They *await their* leader.
 19 *Either* make another *effort* or return.
 52 The *famous estate* is close.
 16 In this *section* little *progress* was made.
 53 The *measure* is *due* to pass.
 16 A *position* in the *field* is his.
 42 To *whom* was the *command* given?
 8 *Whose claim* was bought?
 29 He *represents* the *firm* in this matter.
 2 Go *forward* in that *direction* to reach the city.

When the second hand reaches 43, stop the writing.

Allow no corrections or additions to be made. Ask the pupils to turn their papers over and write their name and grade. Appoint two or three pupils to collect the papers.

TABLE XXVII. MEDIAN SCORES (WORDS SPELLED CORRECTLY) AND AYRES' NORMS. TOTAL WORDS IN EACH TEST IS FIFTY

	GRADE									
	III	IV	V	VI	VII	VIII	IX	X	XI	XII
1917 TESTING										
Number of pupils	1342	1539	1427	1338	1136	876	289	251	188	96
Median (May)	28.6	39.4	32.8	41.8	35.1	43.0	43.3	45.1	46.5	47.9
1919-20 TESTING										
Number of pupils	437	423	440	302	218	153				
Median	28.2	43.0	39.7	43.0	32.6	38.8				
Ayres Norms	33.0	42.0	36.5	42.0	39.0	44.0				

In Table XXVIII the distributions of the scores are given for grades III to VIII inclusive. In the fourth, sixth, and eighth grades there are a number of perfect scores which show that the test was not sufficiently difficult for the best spellers in these grades and, therefore, did not give us a measure of their spelling ability. This statement probably applies also to those pupils who spelled incorrectly one, two, or three words. From the standpoint of accurate measurement of the spelling ability of "good spellers," these tests are too easy, especially for grades four, six, and eight and the high school.

This table is also significant in another respect. The range in the number of words spelled correctly extends from zero, or at the most two, up to fifty and, as has been pointed out in the above paragraph, the spelling ability of some pupils probably extends beyond perfect scores on these tests. That the

TABLE XXVIII. DISTRIBUTION OF SPELLING
SCORES ON MONROE'S TIMED SENTENCE
SPELLING TESTS

SCORE	GRADE					
	III	IV	V	VI	VII	VIII
50	7	45	13	40	14	33
49	11	53	23	83	21	48
48	13	82	29	75	31	64
47	17	76	44	81	45	60
46	24	83	30	82	40	55
45	28	89	38	77	38	61
44	17	63	50	78	39	59
43	34	64	38	81	54	59
42	36	58	49	60	34	48
41	37	45	49	62	30	40
40	35	76	55	69	37	33
39	39	57	31	47	40	25
38	29	67	35	48	34	29
37	30	42	55	43	30	28
36	43	53	44	34	44	17
35	45	43	42	38	40	25
34	43	52	41	23	29	22
33	27	36	40	32	40	18
32	30	43	46	20	31	19
31	42	30	40	24	41	12
30	35	33	46	33	29	24
29	35	33	41	16	23	13
28	31	34	35	17	23	12
27	32	44	29	16	23	10
26	30	23	40	13	28	7
25	34	16	39	18	32	14
24	37	23	42	9	23	3
23	38	26	27	15	13	10
22	26	19	16	3	27	5
21	30	21	30	13	14	5
20	39	10	25	11	18	4
19	23	9	21	5	24	4
18	22	6	24	7	14	1
17	21	6	16	9	18	2
16	30	9	18	9	21	1
15	24	13	22	6	20	
14	26	8	17	7	10	1
13	23	8	14	5	7	2
12	21	5	18	5	10	
11	22	2	15	4	8	1
10	23	8	9	2	8	
9	23	6	16	4	5	
8	25	1	9	2	4	
7	18	2	10	3	5	1
6	12	3	7	1	2	
5	10		6	3	6	
4	16		12		2	
3	11	4	11	2	4	
2	11	4	6	2	1	1
1	12	3	10	1	1	
0	15	3	4		1	
Total	1342	1539	1427	1338	1136	876
Median	28.6	39.4	32.8	41.8	35.1	43.0

pupils of a grade exhibit a wide range of ability is a well known fact. We also know that there is a very large degree of over-lapping between successive grades. Probably all would admit that a grouping of pupils for purposes of instruction exhibiting such great differences in spelling ability does not make for the most effective instruction. Correction of this condition is sometimes difficult. In ungraded schools or in rooms where there are two or more grades, it is relatively easy to group the children for spelling instruction, but in rooms where all the pupils belong to the same grade this is more difficult. The pupils who make perfect scores or scores nearly perfect might be excused from spelling but this would not take care of those most needing instruction.

In view of the fact that the time devoted to spelling is more frequently used in *testing* spelling than in *teaching* spelling, this table suggests that there is a definite need for teaching some pupils to spell. Some pupils appear to have learned to spell under the type of "instruction" which the schools now provide, but a large number of pupils have not done so. These pupils require a different type of "instruction." They probably need some assistance in learning to spell.

Spelling errors. In the 1917 testing, several cities returned, in addition to the class record sheets on which were recorded the scores, the test papers. The specific errors in spelling the test words have been carefully tabulated for 430 third grade pupils, 463 fourth grade pupils, 294 fifth grade pupils, 188 ninth grade pupils, 120 tenth grade pupils, 107 eleventh grade pupils, and 169 twelfth grade pupils. This study of errors was not extended as originally planned to the pupils in the sixth, seventh, and eighth grades because of a lack of funds.

The fundamental reason for giving tests is to secure information concerning the abilities of pupils. In order that this information may be most helpful to us it is necessary that we interpret the scores in terms of pupil needs. The fact that a pupil has a certain score and that this score is above standard or below standard becomes significant to the teacher only when this fact is expressed in terms of the learning needs of this pupil. In the case of scores below standards we have only partially interpreted them when we say that the pupils having such scores need to place more emphasis upon their spelling or that they need assistance from the teacher. We carry the interpretation a step further when we specify what words the pupils need to give their attention to. We may carry our interpretation another step by ascertaining the particular errors which the pupils are making and hence need to correct.

The tabulations of the misspellings show that some ways of misspelling a word occur much more frequently than others. In general about 60 percent of the misspellings belong to one of five kinds and about 30 percent are included in one. If a teacher knows the types of misspellings which are most likely to

found in the examination of 463 fourth grade test papers. Some of the misspellings which occur only once or at the most two or three times probably should not be counted as "true" misspellings. The writing of c-o-l probably should be counted as an incomplete spelling. The writing of c-h-e-r-d indicates either that the pupil is not acquainted with the word "collect" or was unable to recall automatically the spelling, or failed to understand the word in the teacher's dictation. On the other hand the misspellings which recur frequently may be considered as "true" misspellings and are the ones which should be guarded against in the teaching of spelling.

In interpreting spelling scores, we need to consider the questions, "Are all misspellings to be treated alike? Does one misspelling mean the same

TABLE XXX. SHOWING MOST FREQUENT MISSPELLINGS OF THE TEST WORDS
IN GRADE III BY 430 PUPILS

CORRECT FORM OF WORD	No. OF MIS- SPELLINGS	FIVE MOST FREQUENT MISSPELLINGS									
		1		2		3		4		5	
account	218	account	63	count	46	ocount	10	mony	8	aecont	5
again	177	agin	52	agan	40	agen	30	agian	14	agun	3
almost	157	allmost	101	alnost	15	allnost	12	allmast	3	amost	2
anyway	105	enyway	27	eneyway	7	inway	5	annyway	4	inyway	4
army	115	armey	19	arny	15	arme	15	armer	5	armay	4
begin	169	begin	48	begain	36	begun	22	begen	16	begine	6
begin	196	began	49	begin	37	begain	30	begone	9	begon	9
beside	89	by	29	besid	20	becide	5	besaid	3	beson	2
both	72	those	21	bout	6	bothe	3	boath	3	besied	3
bought	185	bot	85	bout	15	brought	13	bough	11	bot	6
change	165	chang	23	changs	14	chain	11	chanes	10	boght	6
children	135	childern	26	children	12	children	5	childred	5	chanse	3
collect	219	colect	36	clect	26	clact	4	clet	4	childer	3
contract	172	contract	49	contrach	5	conrect	4	contracted	4	clack	4
deal	167	deed	45	dill	28	dell	27	dile	17	contrak	4
died	120	dide	50	did	18	dided	13	dead	7	del	6
dress	55	drees	12	dreess	8	drass	8	drea	2	diad	2
drill	69	dril	21	drell	7	dill	5	grill	4	dreass	3
driven	77	drive	11	driving	5	dreven	5	drivin	5	drazil	4
enter	159	inter	71	anter	13	enture	7	intre	4	drivan	4
extra	174	extry	39	exter	10	antra	5	extery	4	ender	4
few	134	you	48	fue	32	to	8	flew	6	extray	4
follow	123	folow	14	foller	8	fallow	7	follo	6	do	4
goes	162	gose	48	gos	47	go	44	goese	2	flow	6
great	188	grate	91	grat	33	grat	10	grait	8	goe	2
income	140	incom	12	incone	8	come	5	encome	2	grad	3
inform	107	inform	17	reform	12	infrom	9	inform	6	incume	1
members	122	members	18	meners	9	mambers	8	mebers	5	informe	6
might	155	mite	78	mit	24	mint	11	night	5	minbers	4
money	126	mony	63	many	9	maney	8	noney	3	mighe	3
month	149	mounth	42	moth	30	mouth	17	mound	9	momy	3
office	138	ofice	23	ofic	21	offes	16	offer	8	mont	7
omit	188	omet	25	omite	21	oment	15	onit	14	ofs	7
paid	167	pade	51	payed	32	pad	28	pay	18	omeat	11
past	119	passed	30	pass	22	pas	9	pasted	7	payd	5
picture	131	pitcher	17	picher	13	pictur	9	pictors	7	pask	4
please	165	pleas	77	plese	17	ples	8	pleace	7	peteure	7
provide	128	proviad	10	provid	8	proved	8	proveid	7	place	6
railroad	111	railrode	12	rayroad	7	ralroad	4	railroat	3	prove	5
ready	144	redy	36	rady	22	read	9	redey	7	rollroad	3
recover	120	cover	22	uncover	20	recove	9	recuver	8	rade	5
return	98	reture	13	returne	7	retun	5	returne	4	recome	3
says	138	sais	19	said	17	say	13	said	10	retern	9
those	140	thoes	76	thos	14	thouse	6	thous	5	sas	4
ticket	69	tick	18	ticked	5	tickit	3	ticke	3	thoses	3
took	88	tuck	16	tock	13	take	7	tooke	5	tecket	5
unable	82	unabel	10	onabel	6	unabil	5	unble	5	tuk	5
understand	83	unstand	13	under	5	understant	5	undestand	4	onable	3
while	109	whil	22	wile	21	whill	11	will	10	understand	6
who	124	how	48	ho	35	hoe	8	hew	6	why	2
TOTAL	6743		1845		902		544		331		221
PERCENT			27		13		8		5		3

as any other misspelling?" Does "colect" have the same meaning with reference to the pupil's need for instruction as "connect" or "cole" or "ceaclect?" Probably all misspellings do not indicate the same instructional needs. The writing of "colect," as was done by nearly one-fourth of all these fourth grade pupils, indicates uncertainty concerning the double consonant or a wrong habit formed. "Connect" probably indicates that the word was misunderstood. "Ceaclect" indicates the lack of a fixed habit and possibly that the pupil was not acquainted with the word.

In Tables XXX, XXXI, and XXXII, we give a list of the five most frequent misspellings¹¹ of the test words in grades III, IV and V. The column

TABLE XXXI. THE MOST FREQUENT MISSPELLINGS OF THE TEST WORDS IN GRADE IV BY 463 PUPILS

CORRECT FORM OF WORD	NO. OF MIS- SPELLINGS	FIVEMOST FREQUENT MISSPELLINGS									
		1		2		3		4		5	
account	246	account	104	count	42	ocount	12	accout	7	acount	7
again	101	agin	40	agan	27	agian	5	agen	9	agine	2
almost	136	ailmost	80	alnost	28	most	8	ainose	2	almos	1
anyway	76	enyway	21	eneyway	7	any	6	inyway	5	anway	2
army	78	arney	17	arny	7	arm	7	arme	4	arney	4
begin	137	began	52	begain	33	begen	16	begun	10	begon	2
begun	200	began	52	begin	33	begain	25	begon	10	becine	8
beside	55	besid	21	by	9	becide	5	besde	2	aside	2
both	77	bouth	16	those	14	bough	10	bothe	6	bought	5
change	144	chang	21	changes	11	changed	8	chanes	8	chane	8
children	63	chilrdren	11	chilrden	6	chirlden	5	chirldren	4	chidren	3
collect	257	colect	110	clect	21	cleck	9	ciet	5	celect	4
contract	173	conctrack	48	comrtract	25	contrat	8	concrat	7	concrach	6
deal	106	deel	26	dell	24	del	6	dill	6	dile	4
died	72	dide	29	die	9	dided	6	did	6	Died	3
dress	31	drees	7	dreess	3	dres	3	Dress	2	drese	2
drill	59	dril	10	drell	7	dill	4	trill	3	Drill	2
driven	66	drivin	9	dreven	4	drive	3	drivvin	3	driver	3
enter	99	inter	61	enten	3	entere	3	anter	3	ender	3
extra	159	extrly	45	extray	8	extery	6	exter	5	axtra	4
few	132	you	56	fue	32	fiew	5	tue	3	fhew	2
follow	97	fallow	19	flow	11	folow	10	fllo	6	foller	5
goes	88	go	24	gose	30	gos	17	gois	2	was	2
great	95	grate	62	geart	5	grat	4	greate	4	gred	3
income	90	incone	18	incom	16	encome	7	incon	4	icome	3
inform	132	inform	33	imform	16	informe	13	enform	7	enform	6
members	129	members	31	members	7	memers	6	meners	5	membes	3
might	121	mite	57	mit	14	mint	14	migh	5	night	4
money	79	mony	46	noney	7	nony	2	momey	2	moeny	2
month	114	mounth	31	mouth	18	nouth	11	moth	9	mount	5
office	94	ofice	20	ofic	22	omet	19	omitt	10	omeat	10
omit	205	omite	35	oment	15	opat	12	paied	6	Paied	2
paid	103	payed	55	pade	18	pas	5	pasted	4	pass	2
past	69	passed	20	pass	10	Picture	8	pictur	3	pictor	3
picture	75	pitcher	16	pitcure	10	plase	5	plise	5	pleace	4
please	121	pleas	66	plese	14	pervide	8	proved	5	porvide	4
provide	160	provid	19	provid	13	railrad	3	roillroad	2	Rail Road	2
railroad	73	railrode	21	ralrode	8	reddy	6	read	3	raid	2
ready	78	redy	27	rady	17	recove	12	recuver	6	rncover	3
recover	118	cover	40	uncover	8	retern	8	retrun	4	retune	3
return	62	returne	11	retun	8	say	4	ses	6	said	6
says	105	said	24	sais	16	thouse	4	thoes	3	thoses	3
those	96	thoes	58	thos	10	tictet	3	tictet	2	tickt	2
ticket	46	tick	5	tichet	8	toke	7	tuke	2	take	2
took	47	tuck	16	tock	8	unabel	5	inable	5	uable	5
unable	73	anable	16	onable	8	undersand	4	understan	3	understand	3
understand	61	unstand	15	understad	5	whill	7	wile	5	whal	5
while	61	whil	16	will	8	Hew	5	wow	3	whow	2
who	69	how	26	ho	22						
TOTAL	5133		1633		714		388		236		178
PER CENT			33		14		7		5		3

headed "number of misspellings" gives the total number of misspellings. In the following columns the five most frequent misspellings are given with their frequencies.

The column giving the total number of misspellings shows clearly that these words were not equal in difficulty for these pupils, although they were taken from the same columns of Ayres' Scale. This condition indicates a limitation of that scale. This may be due to the fact that in deriving the scale Ayres used scores from many different states and hence the evaluation of the words was more general. It may also be due to the use of the scale as a minimum essential list. A third reason may be the nature of the test which was used.

TABLE XXXII. THE MOST FREQUENT MISSPELLINGS OF TEST WORDS IN GRADE V, BY 294 PUPILS

CORRECT FORM OF WORD	No. of misspellings	FIVE MOST FREQUENT MISSPELLINGS									
		1		2		3		4		5	
action	35	actin	3	acienn	3	axion	1	actshon	1	axchian	1
addition	127	adddion	24	adition	21	addadion	8	addition	8	additton	8
attempt	144	atempt	34	attemp	12	attemp	11	atempt	6	attemp	6
await	136	awake	62	awate	18	awaked	15	wait	8	awaite	4
claim	145	clame	79	plain	12	clain	12	clain	11	claime	6
command	111	comand	62	comman	5	commend	4	commad	4	comend	4
connection	95	conection	47	conecton	4	coniction	3	conections	2	connect	2
declare	152	declared	65	declair	17	declaired	7	declar	6	declard	5
due	78	dew	35	do	29	doe	3	du	2	doo	2
effort	122	efert	28	efter	21	efort	15	effor	6	effor	5
Either	155	Eather	99	Ether	20	Eathe	3	either	2	Eaither	2
employs	178	employ	43	employes	17	inplys	13	implyes	10	employ	7
entire	123	intire	80	inter	7	intier	6	intiar	3	intre	2
estate	99	astate	41	state	23	asstate	4	estait	2	states	2
direction	84	dirrection	14	drection	10	way	7	derection	5	dricction	3
famous	109	famious	27	famis	10	famos	7	fames	5	famest	5
field	33	feild	22	fields	2	feeld	1	feilds	1	fielded	1
firm	69	ferm	17	firn	11	furm	9	fern	8	form	3
forward	74	foward	28	foreward	14	frward	3	ford	1	forwar	1
factory	72	factor	13	facture	8	factry	4	fatory	3	factory	3
final	88	finel	34	finial	11	finally	5	fineal	3	finall	3
gentlemen	113	gentleman	41	Gentlemen	9	gentle	5	gentlmen	4	gentlenen	3
important	117	important	18	inportant	7	inportent	7	importane	6	imporant	6
information	129	imformation	32	informaton	7	message	5	imfermaton	5	informaton	3
measure	55	masure	24	mesure	3	maisure	2	meashure	1	maisure	1
present	57	presant	13	preasant	3	preasant	3	prisent	2	presint	2
president	131	President	33	presedent	9	President	6	Presedent	4	presedient	3
Prompt	123	Prompt	40	prompt	13	Prompt	7	Prompted	6	pront	4
property	89	propty	24	proptery	9	propity	6	propety	5	propery	4
progress	101	progus	9	progress	8	progres	8	prograss	5	progrees	3
possession	289	possion	42	postion	34	posion	25	physician	9	possession	7
region	137	regin	20	regon	20	reigon	12	reagon	12	regen	7
Remember	156	Rember	98	rember	12	Remeber	8	Remmber	6	remember	4
represents	132	represents	16	repersents	13	represent	13	reperzent	4	reperzents	3
result	30	rezult	4	reselt	3	ruselt	2	resuled	2	reit	2
Saturday	123	saturday	37	Saterday	27	Sat.	5	Saturdy	5	Saturday	5
section	54	secton	9	seccion	6	sexion	2	seccion	2	Section	1
statement	97	statment	41	stament	5	statemetal	4	satment	3	statemen	3
special	104	specil	11	spechal	11	spechal	3	speshel	3	speache	3
Suppose	188	Suppose	67	suppose	56	Sopose	19	Soppose	9	supose	3
surprise	183	suprise	108	supprise	26	suprize	7	surprize	6	supprize	3
Sometimes	46	Sometime	13	Sometimes	6	Sometime	5	Sometime	3	Sontimes	3
whom	96	shome	17	hom	14	who	13	who	9	hoom	9
Whose	127	Whos	31	Sho's	23	Who	21	Whoes	16	whos	7
women	83	Wemon	24	wemen	13	woman	11	weman	10	wimen	4
wonderful	86	wounderful	23	wonder	8	wondful	8	woundful	15	wonderfull	5
TOTAL	5075		1642		619		346		239		173
PER CENT			32		12		7		5		3

¹¹This includes in certain cases, the wrong use of the capital letter. A test word beginning a sentence was counted wrong if not capitalized. A test word which should be begun with a capital letter was counted wrong if the pupil failed to do this.

APPENDIX

In the following tables, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90 and 95 percentile scores are given for several tests. In general these are tests which have been widely used and which appear to have some permanence. The interpretation of a percentile score is similar to that of a median. In fact the 50-percentile score is the median. The 20-percentile score is the score above which there are eighty percent of the scores and below which there are twenty per cent. These tables are to be used for determining the position which a pupil occupies in the total distribution of his grade.

TABLE I. MONROE'S STANDARDIZED REASONING TESTS IN ARITHMETIC
FORM I. PERCENTILE SCORES BASED ON APRIL TESTING

CORRECT ANSWER					
Percentile	GRADE				
	IV	V	VI	VII	VIII
95	16.5	21.1	20.0	23.5	18.4
90	14.4	19.1	17.8	21.2	16.5
80	11.7	16.5	14.9	18.5	14.0
70	9.9	14.7	13.2	16.5	12.1
60	8.4	13.0	11.7	14.9	10.4
50	7.0	11.3	10.4	13.4	9.0
40	5.9	9.7	9.0	11.9	7.5
30	4.7	8.0	7.6	10.4	5.9
20	3.5	6.2	6.0	8.5	4.2
10	2.0	4.0	4.0	6.0	2.4
5	1.1	2.4	2.8	4.0	1.4
No. of pupils	2968	2996	3518	2803	2515
CORRECT PRINCIPLE					
95	26.7	33.8	26.5	28.9	29.6
90	23.0	31.5	23.8	27.9	26.6
80	18.5	27.5	20.7	25.5	24.0
70	15.6	25.7	18.2	23.2	21.8
60	13.3	22.2	16.0	21.4	19.5
50	11.3	19.2	14.2	19.7	17.2
40	9.2	16.4	12.7	17.7	15.1
30	7.2	13.7	11.1	15.4	12.7
20	5.2	10.3	9.2	13.0	9.9
10	2.8	6.4	6.6	9.9	6.4
5	1.1	3.9	4.7	7.5	4.1
No. of pupils	2932	3027	3498	2706	2472
RATE					
95	16.4	23.4	18.1	20.4	17.3
90	14.5	19.5	15.5	18.3	14.9
80	11.8	15.9	12.9	15.4	12.0
70	10.0	14.0	11.4	13.7	10.1
60	8.7	12.3	10.0	12.3	8.6
50	7.8	11.2	8.7	11.2	7.5
40	6.7	9.9	7.9	9.9	6.6
30	5.7	8.6	7.0	8.6	5.7
20	4.5	7.4	5.8	7.6	4.7
10	3.0	5.4	4.2	6.1	3.2
5	1.7	3.8	3.0	5.1	1.8
No. of pupils	1412	1705	1699	1717	1642

TABLE II-A. MONROE'S STANDARDIZED SILENT READING TEST. COMPREHENSION. PERCENTILE SCORES BASED ON MAY TESTING.

FORM 1

PERCENTILE	GRADE									
	III	IV	V	VI	VII	VIII	IX	X	XI	XII
95	20.8	24.2	33.8	38.6	39.4	43.9	51.5	50.1	51.8	58.4
90	14.4	20.0	26.7	34.0	35.7	40.4	43.3	45.0	42.1	51.7
80	11.2	16.5	22.6	28.3	30.1	35.4	36.3	39.0	40.0	42.5
70	8.9	14.0	19.4	24.2	26.5	31.6	31.0	33.9	35.9	40.2
60	7.0	12.2	17.1	20.2	23.6	27.5	27.3	29.8	32.1	36.5
50	5.2	10.5	14.9	18.8	21.0	24.1	24.0	25.8	28.5	33.3
40	4.1	8.7	13.2	16.7	18.8	23.6	20.8	21.4	26.0	30.5
30	3.1	7.0	11.4	14.5	16.8	19.4	17.6	19.0	23.2	25.1
20	2.0	5.1	9.4	12.0	14.6	16.7	14.3	16.1	20.2	21.7
10	1.0	2.6	6.2	9.0	11.4	13.1	10.7	10.3	15.3	16.1
5	.5	1.3	4.0	6.1	9.4	10.8	7.4	8.2	11.6	11.3
Number of pupils	1464	1805	1729	1871	1270	1337	783	397	236	230

FORM 2

PERCENTILE	GRADE									
	III	IV	V	VI	VII	VIII	IX	X	XI	XII
95	21.0	27.2	29.7	39.6	43.8	45.8	48.6	52.4	56.3	58.9
90	17.6	24.1	27.2	36.0	40.0	42.9	41.7	46.1	49.7	51.8
80	13.6	20.3	24.6	31.1	35.4	38.5	34.9	39.1	42.9	43.6
70	11.0	17.5	22.3	28.1	32.2	35.2	30.8	34.0	38.0	39.7
60	9.2	14.9	20.0	25.6	29.3	32.6	27.5	30.1	33.1	35.9
50	7.9	13.4	17.9	22.7	27.0	29.9	24.5	26.7	29.4	32.1
40	6.4	12.0	15.7	19.7	24.6	27.6	21.5	23.5	25.9	28.4
30	5.0	10.5	13.5	16.9	21.4	25.3	18.2	20.6	22.8	24.9
20	3.3	8.5	11.6	14.0	17.9	21.6	14.6	16.7	19.9	20.3
10	1.7	6.0	8.0	12.2	14.1	17.1	10.2	12.3	15.1	16.1
5	.8	4.2	6.5	8.0	11.3	14.0	6.8	10.0	11.3	12.6
Number of pupils	8741	10625	10157	9404	8791	7561	781	625	443	320

FORM 3

PERCENTILE	GRADE					
	III	IV	V	VI	VI	VII
95	22.2	25.1	27.7	41.9	44.6	46.0
90	19.4	23.8	25.6	37.7	41.7	43.6
80	15.7	21.3	23.6	31.9	36.9	39.9
70	13.4	19.1	22.2	27.6	32.9	37.0
60	11.5	17.1	20.6	24.8	29.4	35.3
50	9.6	15.2	19.0	22.1	26.5	31.1
40	8.0	13.5	17.3	19.4	23.8	28.3
30	6.4	11.2	15.3	16.6	20.4	25.5
20	4.7	9.0	12.8	13.8	18.2	22.1
10	2.3	5.8	10.1	11.0	13.7	17.5
5	1.2	3.3	7.1	8.5	11.3	13.9
Number of pupils	1604	1680	1704	1580	1513	1186

TABLE II-B. MONROE'S STANDARDIZED SILENT READING TEST. RATE PERCENTILE SCORES BASED ON MAY TESTING

FORM 1

PERCENTILE	GRADE									
	III	IV	V	VI	VII	VIII	IX	X	XI	XII
95	98.0	111.0	143.5	145.4	146.6	148.9	136.6	143.0	138.3	152.1
90	84.5	105.9	126.9	143.8	142.7	146.9	120.9	130.1	130.9	137.4
80	70.7	81.0	110.0	115.5	120.3	142.7	109.5	120.2	121.5	129.4
70	60.6	80.6	96.4	103.6	110.9	135.6	100.8	100.4	105.0	122.7
60	52.2	72.8	87.0	95.3	102.0	115.1	86.7	88.5	92.0	107.6
50	44.5	65.2	78.8	89.0	95.1	105.6	82.6	84.4	86.6	102.7
40	38.0	58.1	73.0	73.0	88.8	97.2	78.1	80.3	83.1	88.2
30	33.0	51.7	65.9	67.6	82.8	91.1	73.3	75.6	79.5	84.2
20	20.6	44.4	56.7	60.2	66.6	82.5	66.3	70.7	74.4	80.2
10	10.6	34.3	44.6	49.8	58.2	62.3	55.3	61.5	67.9	69.2
5	5.2	25.0	34.6	41.1	51.5	53.4	50.0	54.9	58.3	60.8
Number of pupils	1464	1840	1828	1869	1279	1357	785	397	235	220

FORM 2

PERCENTILE	GRADE									
	III	IV	V	VI	VII	VIII	IX	X	XI	XII
95	117.4	126.7	128.4	165.9	167.2	168.0	137.9	138.7	140.4	147.9
90	98.5	122.2	126.3	161.7	164.4	165.9	130.4	132.2	134.1	139.2
80	78.0	107.0	122.1	147.0	157.3	161.8	123.6	123.6	125.8	130.7
70	71.8	96.9	114.3	136.3	146.0	154.6	107.6	107.2	109.8	123.5
60	67.0	90.7	104.6	124.3	137.4	145.2	101.5	103.2	104.1	107.5
50	62.6	76.6	97.2	114.1	131.5	137.4	87.2	88.6	89.0	102.3
40	55.7	71.9	91.5	107.8	115.8	131.5	82.9	83.8	84.3	87.6
30	45.5	67.1	76.9	100.8	109.1	115.5	78.1	79.0	79.6	83.1
20	36.8	62.2	71.3	82.7	101.2	107.7	72.0	74.1	73.1	77.5
10	28.9	49.5	63.1	72.1	80.5	88.1	62.9	67.1	61.9	69.6
5	15.5	37.8	55.4	54.9	72.2	79.7	56.1	57.2	54.7	59.8
Number of pupils	9739	10579	10150	9853	8767	7681	781	626	443	322

FORM 3

PERCENTILE	GRADE					
	III	IV	V	VI	VII	VIII
95	123.3	127.0	128.3	146.7	148.3	148.8
90	109.9	123.9	126.7	142.8	146.2	147.2
80	99.4	114.7	123.3	124.1	141.9	143.9
70	89.0	106.6	119.9	114.9	124.6	140.5
60	84.5	101.6	109.6	107.7	115.9	124.0
50	80.0	93.8	105.6	101.1	109.0	117.3
40	72.3	86.9	101.6	94.1	102.5	112.2
30	55.1	81.4	94.2	86.9	95.0	106.0
20	46.7	72.6	85.7	79.6	86.8	98.3
10	42.7	48.9	75.6	69.1	77.5	85.2
5	22.7	43.8	59.6	59.2	65.3	76.5
Number of pupils	1608	1695	1702	1579	1528	1179

TABLE III-A. CHARTERS' DIAGNOSTIC LANGUAGE TESTS FOR GRADES
III TO VIII. PRONOUNS. FORM I. PERCENTILE SCORES
BASED ON MARCH TESTING.

PERCENTILE	GRADE					
	III	IV	V	VI	VII	VIII
95	27.9	27.7	30.7	32.2	35.4	38.9
90	24.8	24.8	27.5	29.7	33.3	37.5
80	21.3	21.5	23.6	26.9	30.6	35.0
70	18.5	19.3	21.7	24.7	28.4	32.9
60	15.8	16.9	20.1	22.9	26.3	31.1
50	13.6	15.1	18.5	21.4	24.5	29.0
40	11.7	13.5	16.8	19.7	22.5	26.9
30	9.9	11.9	15.2	17.9	20.7	24.4
20	7.8	10.0	13.3	16.0	18.5	21.3
10	5.1	7.6	10.9	13.2	15.5	16.0
5	3.3	5.8	9.5	10.6	13.1	10.9
No. of pupils	787	864	895	1344	1566	1253

TABLE III-B. CHARTERS' DIAGNOSTIC LANGUAGE TEST FOR GRADES
III TO VIII. VERBS (FORMERLY VERBS A.) FORM I. PER-
CENTILE SCORES BASED ON MARCH TESTING.

PERCENTILE	GRADE					
	III	IV	V	VI	VII	VIII
95	29.2	32.0	34.4	37.0	36.3	39.4
90	25.3	29.0	32.2	34.3	34.7	38.4
80	20.5	24.3	29.5	31.0	32.7	36.8
70	17.8	21.6	27.4	28.1	31.1	35.5
60	15.0	19.5	25.4	26.1	29.4	34.2
50	12.6	17.7	22.6	24.3	27.7	32.8
40	10.8	15.9	20.7	22.4	25.6	31.2
30	8.2	13.9	18.6	20.2	23.4	29.6
20	6.2	11.7	15.3	17.7	21.5	27.2
10	3.9	8.8	11.5	13.6	18.6	23.4
5	2.7	6.4	7.4	11.3	16.4	20.8
No. of pupils	365	403	373	478	539	638

TABLE III-C. CHARTERS' DIAGNOSTIC LANGUAGE TEST FOR GRADES
III TO VIII. MISCELLANEOUS A (FORMERLY MISCELLANEOUS)
FORM I. PERCENTILE SCORES BASED ON MARCH TESTING

PERCENTILE	GRADE					
	III	IV	V	VI	VII	VIII
95	27.9	23.8	22.3	29.7	31.9	34.6
90	24.1	18.6	19.1	26.6	28.9	31.6
80	15.9	15.1	16.9	23.0	25.8	28.3
70	11.5	12.6	15.0	20.6	21.9	26.2
60	8.7	10.8	13.3	18.6	19.7	24.1
50	6.7	9.3	11.6	16.5	18.9	22.3
40	5.7	7.9	10.2	14.6	17.2	20.3
30	4.6	6.4	9.0	12.7	15.5	17.9
20	3.3	5.1	7.3	10.8	13.4	15.4
10	2.0	3.4	5.5	8.5	10.4	12.0
5	1.4	2.6	4.1	6.5	7.9	9.1
No. of pupils	386	669	668	845	758	494

TABLE III-D. CHARTERS' DIAGNOSTIC LANGUAGE TEST FOR GRADES
III TO VIII. MISCELLANEOUS B (FORMERLY VERBS B). FORM I.
PERCENTILE SCORES BASED ON MARCH TESTING

PERCENTILE	GRADE					
	III	IV	V	VI	VII	VIII
95	27.5	31.7	33.7	38.7	39.0	38.5
90	20.2	29.5	31.5	36.4	37.2	37.3
80	16.7	25.9	28.8	33.5	34.6	35.7
70	13.3	22.9	26.4	31.4	32.8	34.6
60	10.3	20.2	24.0	29.4	31.2	33.4
50	7.9	17.8	22.0	27.3	29.4	32.0
40	5.0	14.6	19.5	24.7	27.4	30.7
30	3.7	11.9	16.9	21.7	25.4	28.9
20	2.5	9.4	14.4	18.0	21.9	26.7
10	1.3	6.5	10.9	13.9	18.2	24.3
5	0.7	3.7	7.1	10.2	15.5	21.2
No. of pupils	230	430	307	475	412	294

TABLE IV. CHARTER'S DIAGNOSTIC LANGUAGE AND GRAMMAR TEST
FOR GRADES VII AND VIII. FORM I. PERCENTILE SCORES BASED
ON APRIL TESTING

PERCENTILE	Pronouns		Verbs		Miscellaneous	
	GRADE		GRADE		GRADE	
	VII	VIII	VII	VIII	VII	VIII
95	27.9	33.4	36.4	34.5	20.9	30.6
90	23.5	32.5	34.1	31.5	17.0	26.4
80	17.6	27.7	27.4	26.3	12.9	21.7
70	13.0	24.0	17.6	21.3	10.4	17.0
60	10.2	20.5	11.8	16.8	7.8	14.4
50	8.0	17.1	7.8	14.0	6.3	11.9
40	6.3	13.7	5.0	10.9	5.1	9.5
30	4.9	10.2	3.5	8.3	3.3	7.2
20	3.5	7.0	2.3	5.4	2.3	5.1
10	2.1	4.2	1.1	3.0	.9	3.1
5	1.3	3.0	.5	1.9	.5	2.2
No. of pupils	936	657	434	497	332	362

TABLE V-A. WILLING'S SCALE FOR MEASURING WRITTEN COMPOSITION,
STORY VALUE. PERCENTILE SCORES BASED ON MARCH TESTING

PERCENTILE	GRADE					
	IV	V	VI	VII	VIII	IX
95	76.3	89.9	93.2	95.7	96.1	95.3
90	66.1	85.5	88.2	91.5	96.0	90.8
80	57.5	77.7	81.5	88.1	88.6	87.5
70	52.1	71.3	76.8	82.8	84.6	84.9
60	46.8	65.0	72.6	78.0	80.5	82.2
50	41.5	58.7	68.1	74.0	76.6	79.0
40	36.9	52.8	63.3	69.2	72.9	73.8
30	32.7	46.6	58.0	63.5	68.7	68.4
20	28.4	40.2	51.5	50.1	62.2	62.3
10	24.2	31.5	43.1	46.0	54.7	53.3
5	22.1	26.0	37.0	37.4	46.9	43.6
No. of pupils	320	579	705	692	572	129

TABLE V-B. WILLING'S SCALE FOR MEASURING WRITTEN COMPOSITION,
 ERRORS PER 100 WORDS. PERCENTILE SCORES BASED
 ON MARCH TESTING

PERCENTILE	GRADE					
	IV	V	VI	VII	VIII	IX
95	32.1	28.8	18.8	18.1	14.8	11.9
90	31.7	24.9	16.0	13.9	11.9	10.3
80	28.2	19.4	11.9	10.8	8.8	7.9
70	23.9	15.6	9.9	8.7	7.4	6.0
60	20.9	12.5	8.2	7.3	6.1	5.2
50	18.5	10.7	6.8	5.8	5.2	4.4
40	15.5	9.3	5.5	4.8	4.4	3.6
30	13.0	7.3	4.2	3.8	3.6	2.8
20	10.3	5.3	3.0	2.6	2.5	1.9
10	7.3	3.5	1.6	1.3	1.2	.9
5	5.4	2.1	.7	.7	.6	.5
No. of pupils	327	578	702	693	473	129

TABLE VI. HARLAN'S TEST OF INFOR-
 MATION IN AMERICAN HISTORY
 PERCENTILE SCORES BASED
 ON MAY TESTING

PERCENTILE	GRADE	
	VII	VIII
95	78.8	95.2
90	70.9	94.1
80	56.6	92.2
70	54.0	81.1
60	48.8	74.9
50	43.9	68.2
40	38.8	56.7
30	31.9	50.9
20	27.1	41.3
10	20.1	31.4
5	15.5	27.7
Number of pupils	1109	1691

TABLE VII. HOTZ'S FIRST YEAR ALGEBRA SCALES, SERIES A
PERCENTILE SCORES BASED ON MAY TESTING

PERCENTILE	Addition and Subtraction		Multipli- cation and Division		Problems		Equations and Formulas		Graphs	
	GRADE		GRADE		GRADE		GRADE		GRADE	
	IX	X	IX	X	IX	X	IX	X	IX	X
95	11.7	10.5	10.4	10.7	10.6	9.3	12.3	11.0	8.0	7.5
90	10.8	9.7	9.7	10.0	9.9	7.6	11.5	10.4	7.5	6.9
80	9.6	8.9	8.7	9.0	9.0	6.6	10.3	9.5	7.2	6.3
70	8.6	8.4	8.1	8.4	8.1	5.9	9.3	8.8	6.8	5.8
60	7.7	8.0	7.6	7.8	7.3	5.3	8.6	8.2	6.5	5.4
50	6.9	7.3	7.2	7.4	6.4	5.0	7.7	7.9	6.2	5.0
40	6.3	6.7	6.6	6.9	5.7	4.6	7.0	7.5	5.8	4.7
30	5.7	6.1	6.0	6.2	4.8	4.1	6.4	7.0	5.4	4.3
20	4.8	5.5	5.4	5.6	3.8	3.6	5.7	6.4	5.0	3.9
10	3.7	4.3	4.6	4.9	2.6	2.9	4.5	5.5	4.2	3.3
5	2.9	3.2	4.0	4.2	1.5	2.2	3.7	4.3	3.7	2.8
Number of pupils	561	390	570	388	566	394	478	385	121	413

TABLE VIII. HOLLEY'S SENTENCE VOCABULARY SCALES. PERCENTILE
SCORES BASED ON APRIL TESTING

Percentile	GRADE										
	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
95	72.7	53.5	45.2	50.9	65.5	63.9	66.5	68.5	75.2	77.9	82.3
90	68.5	44.7	40.0	47.6	61.0	58.5	62.6	64.6	70.0	75.0	78.2
80	63.8	32.3	35.7	42.4	54.7	52.5	57.4	58.7	66.0	69.6	73.1
70	59.6	25.7	31.6	38.5	49.2	48.2	54.2	55.5	61.8	66.4	68.8
60	57.2	20.0	28.2	35.8	46.2	45.1	50.9	52.3	58.7	63.1	65.8
50	54.8	16.7	25.1	33.0	42.8	41.9	47.8	49.0	56.1	59.9	62.8
40	52.4	13.4	22.1	30.3	39.5	38.4	44.7	45.8	53.4	56.9	59.7
30	50.0	10.1	18.7	27.0	35.5	34.3	41.6	42.6	50.8	53.9	56.3
20	46.1	6.8	14.7	23.6	31.4	30.2	36.8	38.9	46.6	51.0	52.8
10	42.2	3.4	10.7	20.2	24.9	20.9	30.0	33.9	41.6	45.4	48.3
5	40.3	1.7	6.0	13.1	21.1	13.4	22.8	31.3	37.4	42.1	44.1
No. of pupils	117	406	520	465	450	1188	1047	253	223	155	108